9.0 INDICES OF ABUNDANCE

9.1 Introduction

In theory, ecologists and wildlife managers depend on a range of sophisticated methods for assessing population abundance. In practice, it is often necessary to rely on some sort of index to abundance, supplemented perhaps by occasional use of the technically more satisfactory methods on a few sample areas. Almost always the limitation is simply one of costs. If one must deal with large areas it simply is too expensive to use the better methods on a regular basis. Dice (1941:402) expressed the general idea very well:

> "The difficulty of obtaining accurate counts of the number of individual mammals present on a given area has led to attempts to develop indices of abundance for the species concerned. Such indices may or may not be convertible into terms of population density. For many practical uses, however, it is sufficient to know the relative abundance of a particular species in different areas or at different times without having an exact count of the population".

Most of the experience in the 60 years since Dice's statement was published tends to show that exact counts are often not feasible. Estimates of some kind are about all that can be managed, and these are usually difficult to achieve.

One of the risks in using an index is that it may not accurately reflect actual population trends. White-tailed deer (*Odocoileus virginianus*) pellet group counts provide one of the best examples of a useful index of abundance, yet Ryel (1971) found that drive counts on a fenced Michigan area yielded an inverse correlation between counts and index values over 11 years of data. He pointed out various reasons for failure of the pellet group count method used in this instance, and gave a good discussion of ways and means for maintaining quality and integrity of the method.

Three approaches to using indices can be considered : (1) Direct conversion to a census method. The pellet group counts provide one of the better known examples, inasmuch as the counts, under proper circumstances, can be converted directly to an estimate of average numbers of deer present on an area. (2) Calibration through ratio and regression methods, including double sampling. A simple linear relationship between an index and an actual estimate of abundance is used to convert the index to an actual estimate of abundance. (3) Calculation of an improved index or a prediction equation. At times, supplemental information may be used to strengthen an index without converting it to a direct estimate. These aspects are summarized in Fig.9.1.

Several sources of variability need to be considered in appraising an index method. One is stochastic in nature, arising from the variability of the chance fluctuations in the births and deaths that result in change in population size. These changes occur even in the presence of constant birth and death rates, diminishing in importance as the population increases in size. Theoretically, such effects can be neglected if fairly large populations are under study. However, we usually can only study part of large populations, either through sampling the large population or by counts on relatively small sample areas. Stochastic effects may then become important. A second source of variability arises when birth and death rates are influenced by

environmental fluctuations, such as an unusual cold spell or other environmental change during the reproductive season. Sometimes such changes are large enough to be labelled "catastrophic" and the results then are usually dramatic enough to attract attention to specific causes. Less dramatic changes may be difficult to detect. A third souce of variability is that engendered by the sampling or observational process through which data on the population are obtained. Trying to sort out these several sources of fluctuations in numbers may be very difficult and deserves more attention than it usually gets.

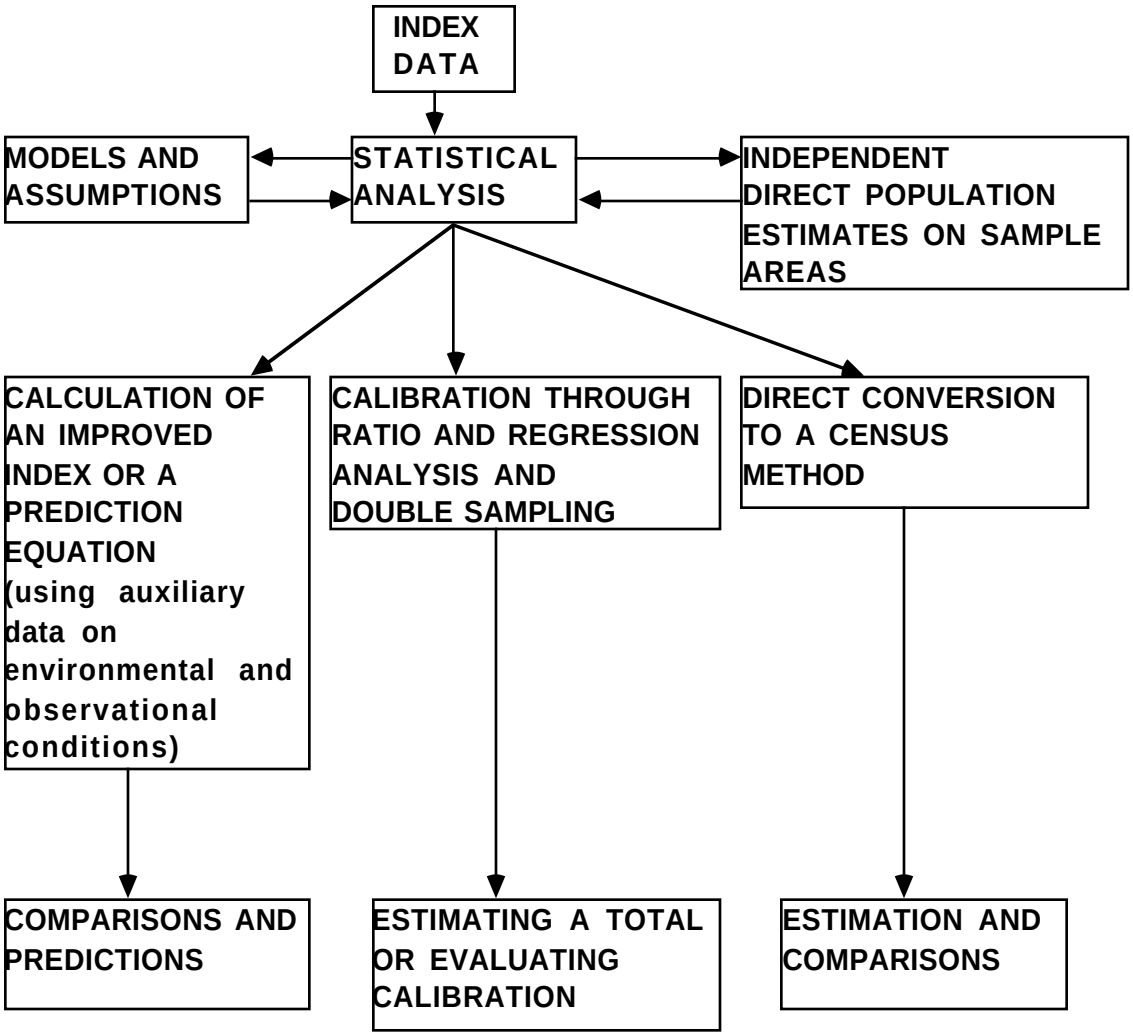

Fig. 9.1 Aspects of the analysis of potential index data.

Indices have received very little attention. The large and extensive reference on estimating animal abundance by Seber (1982) contains only a few pages on indices. This is not Seber's fault, but results simply because there is very little literature on the subject, whereas the stochastic models implicit in tag-recovery and survival data have received much attention in both theory and practice. The presentation here will thus start out by examining a number of sets of data on population trends, and then consider some specific techniques.

Example 9.1 The pellet-group count

One of the better-known wildlife management indices is based on the enumeration of accumulated fecal pellets of ungulates (sometimes politely known as "sign", although the term may also include other kinds of evidence of an animal's presence). Under favorable conditions the method may provide direct estimates of abundance. Such conditions usually depend on a substantial leaf-layer deposited in the fall of the year (or other methods of separating "old" and "new" pellet-groups) plus the over-winter accumulation of groups. Even though many ungulates appear to produce nearly 13 pellet-groups per day (with surprisingly small variance), it is nonetheless true that even an assumption of random distribution of groups still results in the need for rather large numbers of plots being searched to give useful precision in estimation. Thus long accumulation periods are essential. Some workers have resorted to removing pellets from plots (or marking those initially present) in order to be assured of an accurate starting date. Unfortunately, this doubles the fieldwork required.

Although there is a variety of possible sources of error (a major one is simply failure to find all of the groups on a plot), some experience with pellet-group counts on areas where population density could be determined quite accurately by other methods has shown reasonable accuracy in estimating the actual number of deer present. For immediate purposes here, though, we will consider the method as an index. The model is simply:

$$E(x_i) = \beta D_i$$

or that the expected number of pellet-groups per unit area (x_i) is directly proportional to the density (D_i) of animals present on the i^{th} area. The proportionality constant (β) depends on the defecation rate (essentially 13 groups per day), length of accumulation period, and units of measurement of the plots. One further complication is that, if the accumulation period is long (and includes winter conditions and perhaps a hunting season), then there may be an appreciable mortality over the period represented by the counts, so that the estimate obtained is really for an average density.

If independent direct estimates of deer density are available, calibration may be attempted by using ratio or regression methods to convert pellet-group counts to density estimates, i.e., we use a set of direct population estimates (y_i) and a set of pellet-group counts (x_i) on the same areas and turn the above equation around to as to have:

$$D_i = E(Y_i) = (1/\beta)\ E(X_i)$$

so that an estimate of $1/\beta$ is obtained from the comparisons. However, as remarked above, the pellet-group count can also be converted directly. If the mean number of pellet-groups on $1/50^{th}$ acre plots on the ith area is x_i, then a direct estimate is given by (Eberhardt and Van Etten 1956):

$$D_i = 50(640)x_i/12.7(\text{days since leaf-fall})$$

where D_i denotes the estimated number of deer per square mile and $1/\beta$ is calculated from the several quantities on the right-hand side of the equation. Hence, pellet-group counts can be used in all 3 of the ways suggested in Fig. 9.1; directly, by calibration, or by direct conversion. The latter two methods should give the same value for $1/\beta$, within sampling error.

Pellet-group counts represent an excellent example, perhaps the best known, of an index method susceptible to exact treatment by statistical methods. This is because there is an exactly definable population (the total number of pellet-groups) available on a discrete area. In most cases, there will be a considerable advantage to be gained by using stratified sampling methods, and the costs of travel versus those associated with actual counting on plots are such as to dictate the use of a cluster of plots at each location (the individual plots cannot be very large due to the effects of plot size on counting errors).

Some experience has indicated that the negative binomial distribution provides a good fit to the observed frequency of pellet-groups per plot, and this finding may be useful in the efficient design of new surveys.

Example 9.2 Conversion factors for pellet-group counts

Part of the Michigan experience with pellet-group counts as a census method includes counts on two fenced areas in which deer numbers are supposedly known with some accuracy. Data from one of these areas (Cusino Enclosure) is as follows (Ryel 1971:124):

Year	Mean number of groups per plot	Known number of deer per square mile
1953	9.403	28.8
1954	2.252	25.0
1955	1.778	28.1
1956	2.246	29.3
1958	2.943	15.5
	12.622	120.7

Using a ratio estimator, we obtain: $\hat{b}_1 = \frac{\sum y_i}{\sum x_i}$ =12.622/120.7 = 0.105.

Data for a direct estimate can be used to calculate: $\hat{b}_2$= [211(19.56)]/[50(640)] = 0.089. Here, 211 is the average number of days since leaf-fall, and 19.56 is an average number of pellet-groups per deer, adjusted for the sex and age composition of the known number of deer on the area. The direct conversion factor is the reciprocal of $\hat{b}_2$ or 11.2, which is somewhat higher than the reciprocal of $\hat{b}_1$ which is 9.5. One might thus expect to overestimate true deer density by using pellet-group counts,

if $\hat{b}_1$ and $\hat{b}_2$ are indeed as different as suggested here. There are various other problems in comparisons on this and another such area (George Reserve) and these were described in detail by Ryel (1971).

Example 9.3 Roadside counts

Roadside surveys have been widely used in assessing the numbers of many species, and a useful model for such a survey is that used for pellet-group counts. However, β is not so readily defined in this case, being affected importantly by the behavior of the species, cover (habitat) conditions, along with the weather and various other factors such as time of day (many species are most active in early morning and in the evening). Quite a lot of effort has gone into attempts to standardize roadside surveys by taking counts during specified times and weather conditions. Very likely such standardization techniques are suitable for controlling most factors other than habitat differences. So far, little has been done to try to take into account the effect on visibility generated by different intensities of vegetative cover. One obvious prospect is to attempt to record distances from the observer for each individual animal seen, in the manner of line transects, but a suitable model is needed for use of such a correction.

The net effect of the several uncertainties about roadside and other visual counts (such as aerial surveys) is to make it likely that such methods may be reasonably satisfactory for comparisons from year to year on the same routes, but rather less useful for comparisons of routes in different cover types. Conversion will need to be accomplished by use of ratio or regression methods and independent direct estimates of density on a sample of areas, as there is presently no way to write an equation like that used for pellet-group counts for roadside surveys.

Well-known examples 0f calling counts are counts of crowing pheasants, or of the drumming of ruffed grouse, cooing of doves, etc. The usual technique is to make counts for a fixed period of time (typically 2 to 5 minutes) at each of a number of stations or "stops". Normally the counts are made along roads as a consequence of the need to cover sizable areas. As with roadside counts, time of day, weather, and seasonal effects are important. Often a degree of standardization is achieved by making frequent counts on a single route at different times of the season. If the population on that route is assumed to remain constant, curves of calling intensity against time of day (and time of season) may serve as a reference standard for adjusting the other counts.

Auditory counts bring in the hearing acuity of the individual observer as an important additional variable. A useful model may then be:

$$E(x_{ij}) = \beta r_i^2 D_j$$

where r_i represents the radius within which a call may be heard by an individual observer. Presumably such a radius may also depend on cover and weather conditions (plus interference from other sounds -- traffic noise being usually the main offender), so there

may be a "regional effect" as well as an "observer effect" on the recorded counts.

The following table gives results of hunting success and calling counts for Gamble quail in Arizona.

Hunting success and calling counts for Gambel quail in Arizona (Smith and Gallizioli 1965).

			Oracle Junction		Pinnacle Peak		Cave Creek	
Year	Quail per trip	Call count						
	y	x	y	x	y	x		
1958	3.81	61	3.53	83	--	--		
1959	2.70	24	1.37	10	--	--		
1960	6.40	103	3.74	94	2.96	72		
1961	2.57	25	1.20	22	0.64	8		
1962	6.09	75	2.83	59	2.55	64		
1963	4.84	62	1.70	25	1.82	36		
1964	2.91	41	1.60	15	1.38	26		
Totals	29.32	391	15.97	308		9.35	206	
$\hat{b}_1 = \Sigma Y/\Sigma X =$			.075		.052		.045	
$\hat{b}_2 = \Sigma YX/\Sigma X^2 =$			.070		.044		.043	
Regression slopes			.054		.030		.034	
Reg. intercepts			1.200		.94		.45	

Example 9.4 A pocket-gopher example.

Reid et al.(1966) gave an interesting example in which an index (mounds and earth plugs) of pocket-gopher (Thomomys talpoides) abundance is compared with actual abundance, as established by trapping-out gophers on sizable plots. They concluded that the relationship between number of gophers present and the index was curvilinear, but did so by plotting the number of gophers per unit area against the number of signs. Such a plot does indeed suggest a nonlinear relationship. However, there were 2 areas (Black and Grand Mesas) studied in 3 years (1962 to 1964).

If we assume the trap-outs to measure absolute abundance, without sampling error, and adopt the model of simple proportionality, $E(x_i)=\beta D_i$, then the appropriate plot is of signs (Y) against gophers per acre (X). Also, if the areas and years are plotted separately (or distinguished by individual symbols), then it seems that the apparent curvilinearity may really be due to differences in β between years and/or areas. It is further

evident that variability increases with increasing density of gophers. Hence, we adopt the model: $x_{ij} = b_{ij} D_{ij} e_{ij}$, where x_{ij} denotes number of signs in the i^{th} year and j^{th} area, D_{ij} is actual density of gophers, b_{ij} is the proportionality constant between gophers and sign, and e_{ij} is a proportional "chance" error. Since D_{ij} is assumed to be measured without error (or with negligible error), we can arrange the equation in terms of signs per gopher, i.e., consider:

$$y_{ij} = \frac{x_{ij}}{D_{ij}} = b_{ij} e_{ij}$$

and take logarithms:

$$\log y_{ij} = \log x_{ij} - \log D_{ij} = \log b_{ij} + \log e_{ij}$$

Then the data can be subjected to a simple one-way analysis of variance. The results indicate that there are significant differences and, using a multiple-comparison test (Scheffe 1959) shows that Grand Mesa in 1963 is significantly different from its 1964 value and from Black Mesa in 1962. We are thus fairly confident that the main factor in the apparent curvilinear relationship is really a difference between areas and years, but could not, of course, exclude a behavioral difference as a possibility at low densities. The conclusion here, as with the quail example above, is that it will very likely be necessary to include both spatial and temporal data in any initial efforts to calibrate an index.

Abundance of pocket-gopher sign (mounds and earth plugs on one-acre plots) expressed as logarithm (base e) of signs per gopher (Reid et al. 1966).

Black Mesa			Grand Mesa	
1962	1963	1964	1963	1964
1.91	2.76	2.01	2.03	1.82
2.32	2.23	2.02	1.99	2.06
2.05	2.16	2.18	2.17	1.41
2.02	1.92	2.54	2.24	1.69
1.86	2.36	1.55	2.07	1.54
1.31	2.48	1.40	2.88	1.76
1.58	2.65	1.61	2.18	2.27
1.80	1.88	1.83	2.63	2.12
2.05	1.93	1.43	2.42	2.09
2.04	2.14	2.24	2.54	2.04
1.70		2.20		
1.74				
2.07				
$\bar{x}$ = 1.882	2.253	1.910	2.316	1.880
n_i = 13	10	11	10	10
s^2 = 0.0670	.0947	.1395	.0865	.0780

9.2 Trends in abundance

For at least the larger vertebrates, there are few occasions when data on the full course of the growth of a population are available. Usually only a relatively short segment of the record of population size is at hand. A conceptual model of the overall possible course of events is nonetheless useful for interpreting these shorter segments of data. Ecology textbooks describe long-term population growth by the logistic model. Further details of this model appear in the chapter on population models, but the general shape is illustrated by the trend of an elephant seal population (Fig. 9.2). The logistic curve is characterized by rapid initial population growth that slows down over time, with the curve ultimately approaching an asymptotic level (often denoted as K). The approach to an asymptotic value may be erratic with large year-to-year fluctuations. Slowing-down of the growth rate usually is associated with resource limitations of one sort or another, often food or space. Such restrictions make populations highly susceptible to year-to-year weather fluctuations.

The logistic curve assumes a constant decline through time of the rate of increase. Although data are limited, evidence for large mammals (Eberhardt 1977b, Fowler 1981) suggest that a different model may be appropriate, with the rate of population growth virtually constant over much of the range, and then slowing down sharply as the asymptotic value (K) is approached. A simple exponential growth curve may then be adequate to to describe the initial stages of growth. The essential feature for present purposes, however, is the overall sigmoid shape of the curve -- concave-upwards in the early stages, and concave-downwards as growth slows down.

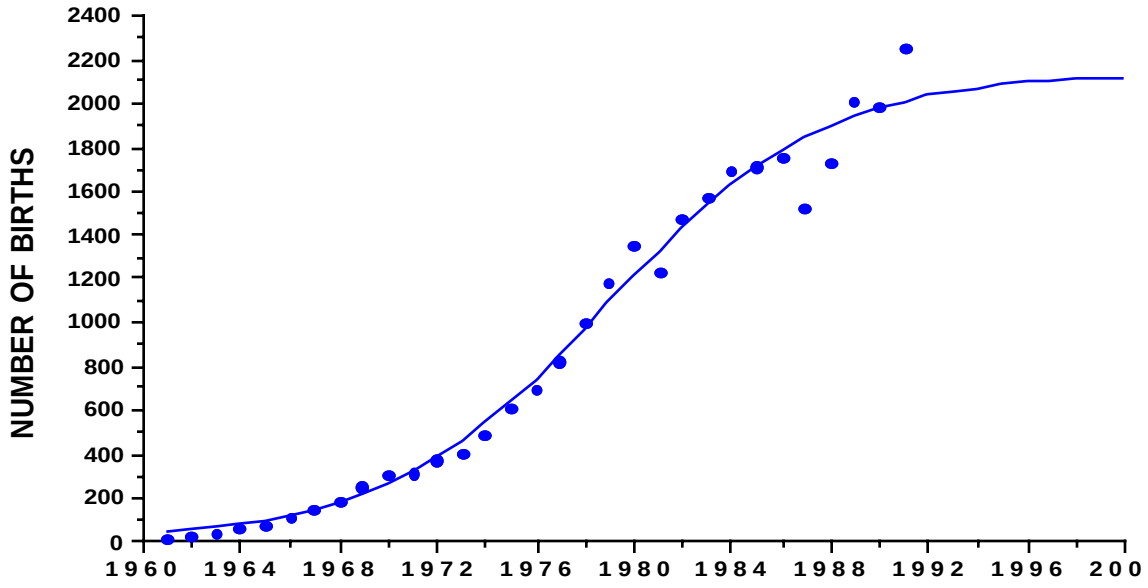

Fig. 9.2. Growth of an elephant seal population (Stewart et al. 1994) with a logistic curve fitted by non-linear least-squares. Data are number of births tallied on colonies on Ano Nuevo Island and the adjacent mainland.

If one accepts generality of the sigmoid curve illustrated in Fig. 9.2, then the analysis of a shorter series of observations will obviously be influenced by the position of the data segment on the overall curve. A basic criterion is simply whether a curve drawn through the data segment is concave-upwards (initial growth stages) or concave-downwards (approaching the asymptotic level). If the early growth segment essentially follows an exponential curve, then plotting logarithms of the counts against time should yield a straight line. Two examples of such data appear in Fig. 9.3.

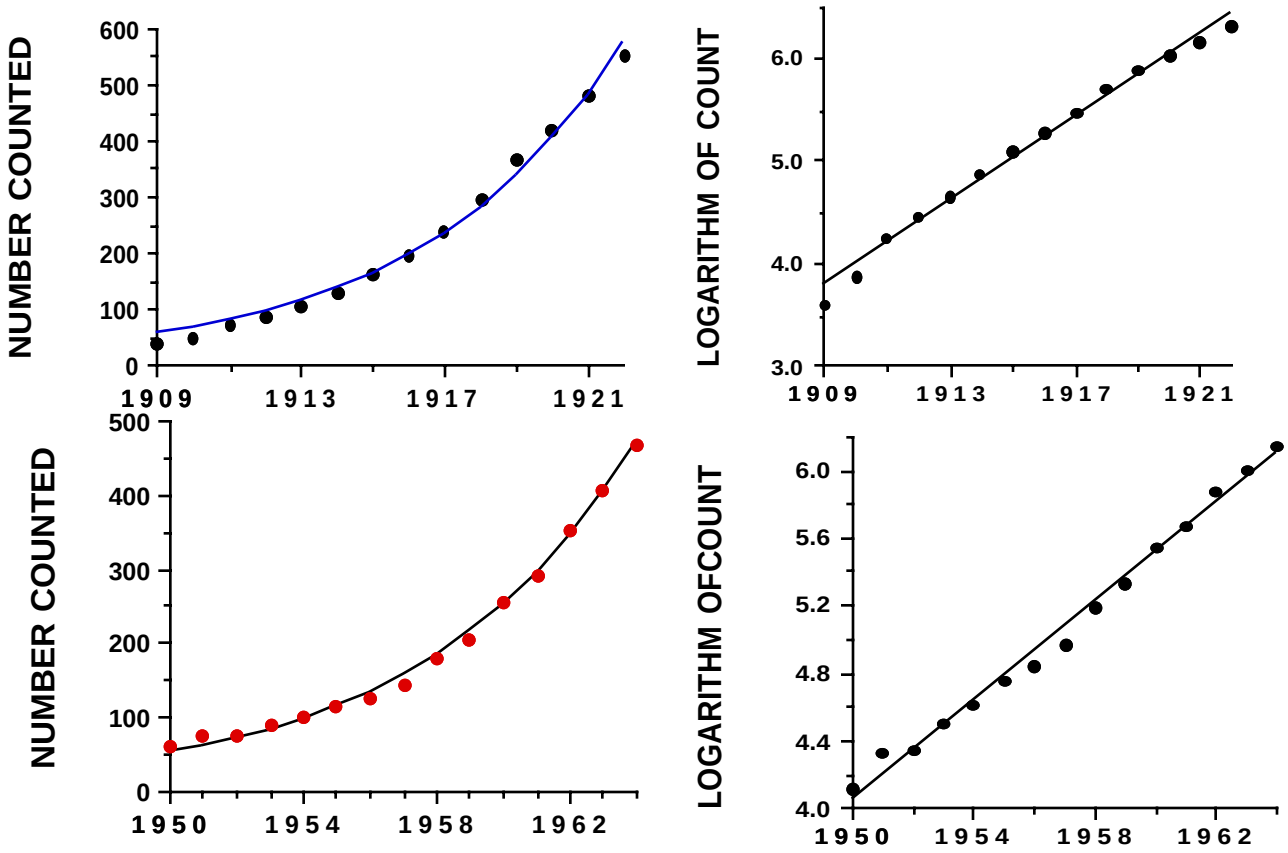

Fig. 9.3. Growth curves for bison (upper panels) with an exponential curve fitted to numbers counted on the left, and a straight line fitted to logarithms of number counted on the right. Data from Fredin (1984). Lower panels show an exponential curve fitted to numbers of Muskox on Nunivak Island, Alaska and the corresponding loglinear regression line. Data from Spencer and Lensink (1970).

Many sets of population trend data are much more variable than those shown above, and we thus need to consider statistical tests. A simple test for curvilinearity is demonstrated in Section 9.3. Various monte carlo simulations were conducted by Eberhardt (1992) to appraise the utility of the approach in attempting to determine whether a segment of population trend data could be demonstrated to be above or below the inflection point on a curve like that of Fig. 9.2 (the inflection point divides a growth curve into 2 segments; the slope of the curve increases up to the inflection point and decreases beyond it). A two-stage test was developed. The observed, untransformed, numbers of the

trend index are checked for curvilinearity in the first stage. If this test is non-significant, then the data are log-transformed and the test again applied. The first stage examines data of the form of the left side of Fig. 9.3. The basis for the second-stage test is the fact that the exponential curve becomes a straight line under log-transformation (right side graphs of Fig. 9.3), so should not show a significant departure from linearity in the test. However, the right-hand limb of Fig. 9.2 remains curved under the log-transform.

9.3 Testing significance of trend lines

In exhibiting curves like those of the left side of Fig. 9.3 above, the exponential curve was used. Growth of large vertebrates may better be described by a closely related curve, the geometric (see Fig. 11.1 and the accompanying discussion). Either curve can be represented by the model of eq.(9.1) by letting $\lambda = e^{rt}$ or $\lambda = (1+r)^t$, as discussed in Chapter 11, so we use the following simple model for a population growing at a constant rate.

$$N_t = N_o \lambda^t \qquad (9.1)$$

where N_t represents abundance at time t, and λ is a measure of the rate of change of the population. Taking natural logarithms converts this model to a linear equation, so we expect the trend to be a straight line if the index used is the logarithm of observed numbers:

$$\log_e N_t = \log_e N_o + t \log_e \lambda \qquad (9.2)$$

We thus tend to plot logarithms of observations of abundance, and to do various kinds of linear regression analyses in exploring the data. In the present section, we will examine a number of sets of data on actual populations and try to infer something about trends from simple regressions.

To study trend, we fit eq.(9.2) by linear regression methods, rewriting it as $y = a + bx$, where $y = \log_e N_t$, $x = t$, and $b = \log_e \lambda$. For purposes of calculation, t will be used here as the sequence of years, 1,2,3,... We are mainly concerned with the slope (b) and variance about regression, so that the observed sequence of dates might serve equally well, i.e., we could use 1967, 1968, 1969, etc. in regression calculations. However, this should not be done in practice due to the prospect of introducing "roundoff" errors in the regression calculations when the x-values are sizable numbers (i.e, one should use the sequence 1,2,3.. for calculations and later plot data against the actual years, 1967, 1968, 1969). A key measure of variability is the "regression mean square", or variance about regression, often written as MS_{reg}. It can be calculated by fitting the linear regression and calculating:

$$MS_{reg} = \{\Sigma\ [y_t - (a + bt)]^2\}/(n - 2) \qquad (9.3)$$

In most instances, it will be desirable to check to see whether there is evidence that the logarithms of the index data appear to change linearly with time, suggesting a constant rate of change in the population. An easy way to do this is to fit a curve to the data, and test for a significant change in the variability around the fitted line. The simplest such curve is the "quadratic" or second degree polynomial. We now fit the function:

$$Y = a + bt + bt^2 \qquad (9.4)$$

and calculate a variance as in eq.(3):

$$MS_{quad} = \{\Sigma\ [y_t - (a + b_1 t + b_2 t^2)]^2\}/(n - 3) \qquad (9.5)$$

In this instance, we use n-3 "degrees of freedom", because another constant (b_2) has been estimated from the data. Actual fitting of eq. (9.4) is readily done by multiple regression. Many computer programs are available to do the fitting, and most of the spreadsheet programs will provide such fits. An F-test of significance can be calculated from the F-ratio:

$$F = \frac{SSLIN - SSQUAD}{MS_{quad}} \qquad (9.6)$$

where $SSLIN = (n-2)MS_{reg}$ and $SSQUAD = (n-3)MS_{quad}$, and significance of the F-ratio is checked in tables of the F-distribution with 1 and n-3 degrees of freedom. These procedures are described in more detail in most introductory statistics texts, and in Chapter 6.

The test of eq.(9.6) can be applied with other alternative curves. For example, a third degree polynomial can also be fitted to the data by multiple regression. The equation is:

$$y_t = a + b_1 t + b_2 t^2 + b_3 t^3 \qquad (9.7)$$

Because this is a more flexible curve than eq.(9.4) it will often appear to give a better fit. However, once a departure from linearity is established, one really needs more information than can be derived from the curve alone in order to assess the situation. Estimates from eq.(9.3) for a number of species appear in Fig. 9.4. Cases where eq.(9.6) gave statistically significant results are denoted by an asterisk (0.05 level of significance) or two asterisks (0.01 level of significance). A very high regression mean (0.250) square for elephant seals results from the nonlinearity evident in Fig. 9.2, and that value is not shown in Fig. 9.4 because of the clear evidence that a different curve is appropriate.

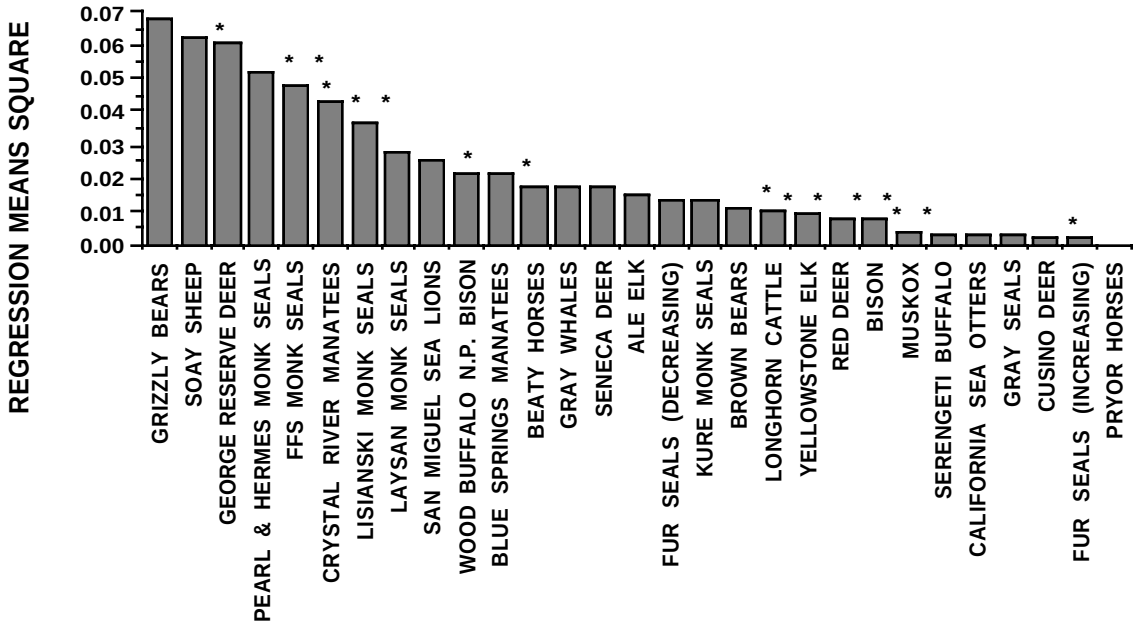

Fig. 9.4. Regression mean square values computed from eq.(9.3) for various species. A single asterisk indicates significance at the 0.05 level for the test of eq. (9.6) and a double asterisk indicates significance at the 0.01 level. Sources and scientific names appear in Table 9.1.

9.4 Assessment of trend data

Fig. 9.4 shows a wide range of regression mean squares, and indicates that the evidence of curvilinearity does not seem to be associated with variability. Some of the likely sources of differences among species can be identified as follows. The mean square for grizzly bears has been assumed to result from the effect of environmental conditions on counts. In wet years, bears can find sufficient food without spending much time in the open and thus are very difficult to find. In especially dry years, they forage widely, and counts tend to be higher in those years. The data used for Fig. 9.4 come from the study reported by Knight et al. (1995). Further analysis of the data yields an improved index and thus a smaller mean square. The variability in the Soay sheep regression evidently results from environmental conditions. Boyd (1974) described the pattern of a build-up for several years, followed by heavy mortality under stress from weather and nutritional conditions. Only 7 years of data are available for the George Reserve deer data set, so it is mainly of interest for the overall observed high rate of increase, and the curvature does not seem to have a recognizable pattern.

Table 9.1 Sources and nature of data and scientific names for data used in Fig. 9.4

NAME	NATURE OF COUNT	SCIENTIFIC NAME	SOURCE
ELEPHANT SEALS	Births	*Mirounga angustirostris*	Stewart et al. 1994
GRIZZLY BEARS	Females with cubs	*Ursus horribilis arctos*	Knight et al. 1995
SOAY SHEEP	Total count	*Ovis sp.*	Boyd 1974
GEORGE RESERVE DEER	Total	*Odocoileus virginianus*	McCullough 1983
MONK SEALS	Beach counts	*Monachus schauinslani*	Gilmartin and Eberhardt 1995
MANATEES	Totals	*Trichechus manatus*	Eberhardt and O'Shea 1995
SEA LIONS	Pup counts	*Zalophus californicus*	DeMaster et al. 1982
WOOD BUFFALO N.P. BISON	Total population estimate	*Bison bison*	Carbyn et al. 1993
FERAL HORSES	Total counts	*Equus caballus*	Eberhardt et al. 1982
GRAY WHALES	Population estimates	*Eschrichtius robustus*	Breiwick 1994
SENECA DEER	Reconstructed pop.	*Odocoileus virginianus*	Hesselton et al. 1965
ALE ELK	Total counts	*Cervus elaphus*	Eberhardt et al. 1996
FUR SEALS (DEC.)	Total counts	*Callorhinus ursinus*	
BROWN BEARS	Spawning stream counts	*Ursus horribilis*	R. A. Sellers, pers. comm.
LONGHORN CATTLE	Total counts	*Bos bos*	Fredin 1984
YELLOWSTONE ELK	Aerial counts	*Cervus elaphus*	Houston 1982
RED DEER	Total counts	*Cervus elaphus*	Clutton-Brock et al. 1982
BISON	Total counts	*Bison bison*	Fredin 1984
MUSKOX	Total counts	*Ovibos moschatus*	Spencer and Lensink 1970
SERENGETI BUFFALO	Population estimates	*Syncerus caffer*	Sinclair 1977
CALIF. SEA OTTERS	Total counts	*Enhydra lutris*	
GRAY SEALS	Births	*Halichoerus grypus*	Bonner 1975
CUSINO DEER	Total	*Odocoileus virginianus*	Ozaga and Verme 1982
FUR SEALS (Inc.)	Total counts	*Callorhinus ursinus*	Kenyon et al. 1954
PRYOR HORSES	Total counts	*Equus caballus*	Garrott and Taylor 1990

The monk seal counts at Pearl and Hermes Reef are highly variable because the seals occupy a number of small islets spread over a sizable area there and thus are difficult to reach and to count. Curvilinearity in the French Frigate Shoals data likely result from incomplete tallies in the early years. Evidence of curvilinearity in the Lisianski monk seal data (Fig. 9.5) seems to be

characteristic of many of the declining populations, including bison in Wood Buffalo National Park, and the decreasing fur seal population. In many such populations, the causes of decline are unknown or imperfectly understood, and likely vary over time, whereas an increasing population often is doing so in consequence of ample food and other resources, and thus is likely to exhibit a smooth pattern.

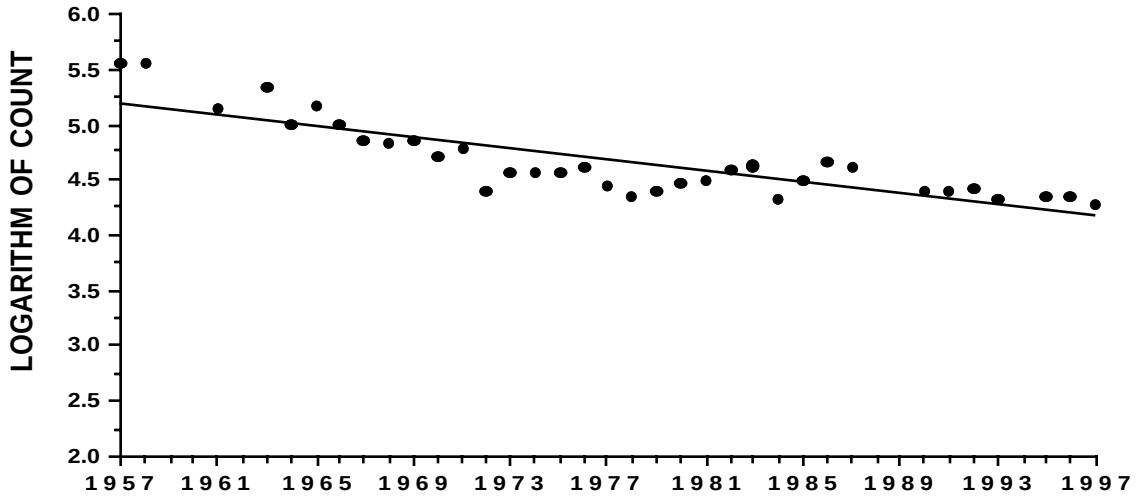

Fig. 9.5. Logarithms of beach counts of monk seals on Lisianski Island.

Somewhat erratic growth patterns may, however, also be evident in increasing populations, as is the case with manatees (Fig. 9.6). Very likely some of the fluctuations result from conditions under which the counts were made, inasmuch as the counts may be made by divers and at times in turbid water (Crystal River).

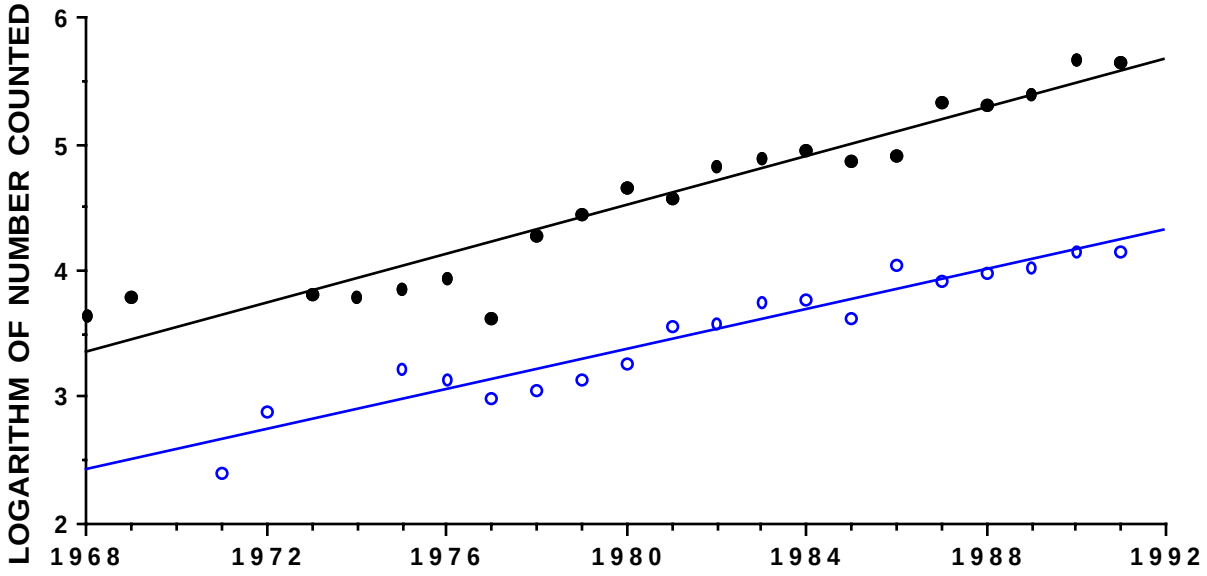

Fig. 9.6. Logarithms of counts of manatees at Crystal River (upper data set) and Blue Springs (lower data points).

It seems quite possible that data on the right side of Fig. 9.4 may approach the circumstance where the fluctuations may largely be associated with the stochastic behavior of the underlying birth and death processes. In several instances (muskox and bison in Fig. 9.3, fur seals (increasing), and Seneca deer), there is a suggestion of an oscillatory pattern in the deviations from linearity. Such oscillations may be a consequence of a changing age structure (Keyfitz 1968 gave the theoretical basis) and are suggested by simulations of longer sequences of observations (Eberhardt 1981). Detection of such an oscillatory pattern is aided by plotting deviations from fitted curves as in Fig. 9.7, which shows the pattern for the muskox data of Fig. 9.3.

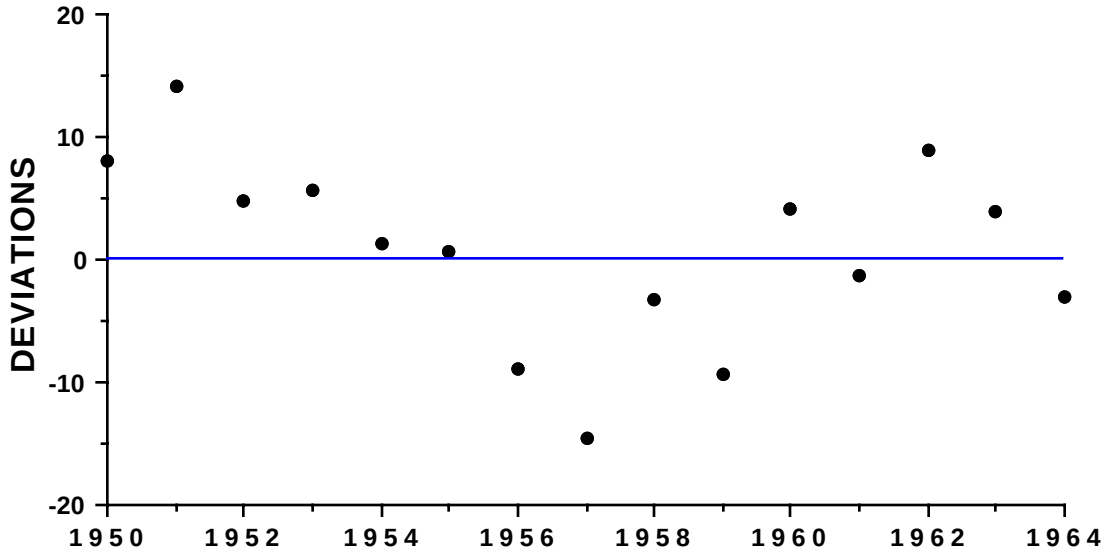

Fig. 9.7. Deviations from the exponential curve fitted to muskox data in Fig. 9.3.

9.5 A test for significant deviations from regression using replicate points.

The test for significant deviations from linearity used in Section 9.4 depends on fitting a curve and testing to see whether the improvement in fit might simply be due to chance. In some cases, replicate counts may be available, so that one can use the variability within years to test significance of deviations from linearity. The advantage here is that we do not need to specify an alternative model like the quadratic or cubic (which may very well be the wrong model). Some counts of brown bears at spawning streams provide an example (Fig. 9.8). In this case, the test consists of making the usual analysis of variance to test for significance of the linear regression (Table 9.2), and then using the pooled variance of individual observations within years to estimate "pure error" (Draper and Smith,1981). The data for calculation of pooled error appear in Table 9.3. A sum of squares of deviations from the mean is calculated for the data in each year where there are two or more observations and these values are summed to give an overall sum of squares, which is subtracted from the "residual" sum of squares in Table 9.2 to yield the "lack of fit" sum of squares (i.e., the variability not accounted for by "pure error"). The degrees of freedom used to calculate pure error (32) is similarly subtracted from the degrees of freedom for residual error to get the degrees of freedom used to calculate a mean square for "lack of fit". An F-ratio as shown in Table 9.2 would be used to test significance of the lack of fit, but

the F-test for regression (1.14) is not significant, so there is no real point in going on to test lack of fit, other than to illustrate the method.

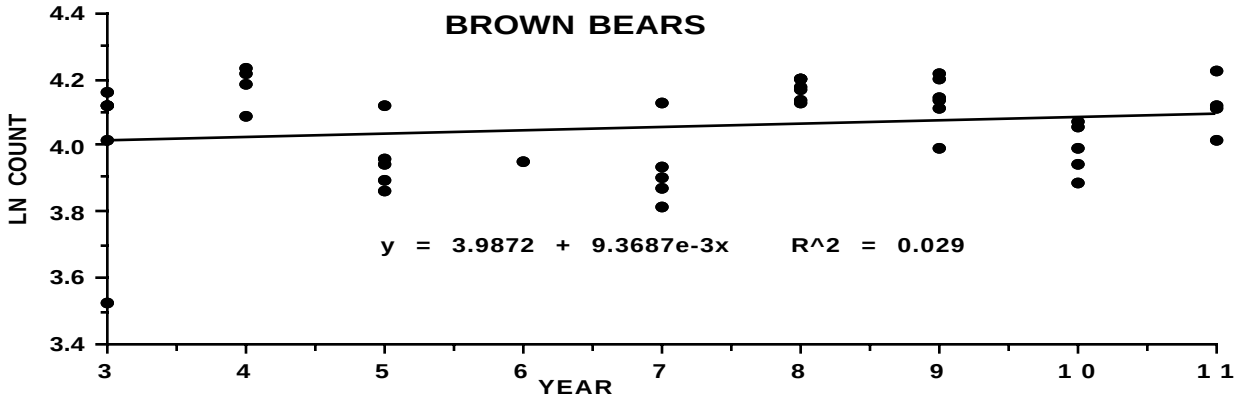

Fig. 9.8 Logarithms of counts of brown bears on salmon spawning streams.

Table 9.2 Test of significance for deviations from regression

	df	*SS*	*MS*	*F*
Regression	1	0.02415	0.02415	1.14655
Residual	39	0.82140	0.02106	
Total	40	0.84555		
Lack of fit	32	.4753	0.01485	
Pure error	7	0.3461	0.04944	

Another example of variability of individual counts is provided by some gray whale data (Breiwick 1994). In this case, replicate counts over time could not be made, as the whales are counted as they migrate past a shoreline counting station. However, estimates of sampling error are available (Table 9.4) and can be used to assess the mean square error obtained in log-linear regression analysis of the data. Because the standard errors of individual estimates were calculated in terms of number of whales, a transformation needs to be used to change to variability on the logarithmic scale. Using the delta method, we find:

$$V(\ln x) = V(x)\left(\frac{1}{x}\right)^2 = \text{Coef. var.}^2(x)$$

Plotting the coefficients of variation against the estimates shows little evidence of correlation, so the average coefficient may be a reasonable estimate of overall variability. The squared value (0.0027) is substantially smaller than the mean square about regression (0.0170) so it seems quite evident that there is a significant departure from linearity, even though the quadratic and cubic mean squares are little different from the mean square from linear regression (Table 9.4). Inspection of the data (Fig. 9.9) suggests the possibility that the population might have reached an equilibrium level. However, the marked decrease in years 4 and 5 suggests the possibility of shifts in the migration pattern.

Table 9.3 Data for computation of "pure error" for brown bear counts

Year	Bears/hour	Logarithm bears/hr	Sum of squares	Deg. of freedom
3	33.85	3.5219		
3	64.04	4.1595		
3	61.88	4.1252		
3	61.2	4.1141		
3	55.24	4.0117	0.2819	4
4	68.7	4.2297		
4	59.3	4.0826		
4	67.9	4.2180		
4	65.3	4.1790	0.0134	3
5	49.4	3.9000		
5	51.4	3.9396		
5	61.6	4.1207		
5	47.4	3.8586		
5	52.45	3.9599	0.0400	4
6	51.88	3.9489		
7	45.14	3.8098		
7	62	4.1271		
7	48.13	3.8739		
7	49.58	3.9036		
7	51.21	3.9359	0.0572	4
8	62.06	4.1281		
8	66.59	4.1986		
8	62.32	4.1323		
8	66.88	4.2029		
8	65.03	4.1748		
8	64.58	4.1679	0.0051	5
9	54.17	3.9921		
9	67.49	4.2120		
9	66.67	4.1998		
9	62.8	4.1400		
9	61	4.1109		
9	62.42	4.1339	0.0311	5
10	48.68	3.8853		
10	51.47	3.9410		
10	58.51	4.0692		
10	57.65	4.0544		
10	54.08	3.9905	0.0238	4
11	61.12	4.1128		
11	55.15	4.0101		
11	68.29	4.2238		
11	61.52	4.1194	0.0229	3
Sums	2386.0800	166.2194	0.4753	32

Table 9.4. Estimate of total numbers of gray whales along with standard errors of estimate and coefficients of variation.

YR	ESTIMATE	S.E.	COEF. VAR.	LN ESTIMATE
1	12921	964	0.075	9.4666
2	12070	594	0.049	9.3985
3	12597	640	0.051	9.4412
4	10707	487	0.045	9.2787
5	9760	524	0.054	9.1860
6	15099	688	0.046	9.6224
7	14696	731	0.050	9.5953
8	12955	659	0.051	9.4692
9	14520	796	0.055	9.5833
10	15304	669	0.044	9.6359
11	16879	1095	0.065	9.7338
12	13104	629	0.048	9.4807
13	16364	832	0.051	9.7028
18	21443	1182	0.055	9.9732
19	20113	927	0.046	9.9091
21	20869	913	0.044	9.9460
26	17674	1029	0.058	9.7798
27	23109	1262	0.055	10.0480

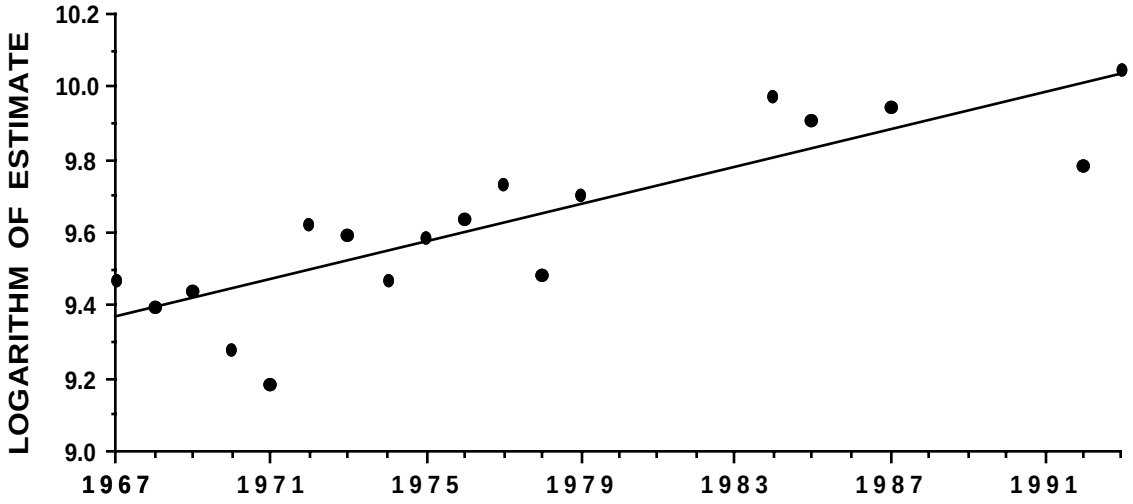

Fig. 9.9 Logarithms of counts of gray whales. Data from Breiwick (1994).

9.6 A test for linearity based on the lowess method

The test for deviations from linearity based on fitting a quadratic curve to the data uses the difference between sum of squared deviations about the loglinear regression line and that about the quadratic curve, and tests this difference against the regression mean square about the quadratic, using an F-test. Such a test can also be used with the cubic equation (or any suitable alternative), with the only change being that of reducing the degrees of freedom to adjust for the extra parameters fitted. An alternative is to consider curves fit with a "locally weighted regression" technique (variously called "loess" or "lowess",

and due to Cleveland 1979). Weighted linear regressions are fit at each point on the graph (e.g., if the data span 30 years, then such regressions are fit at each of the 30 years) by selecting data points in the immediate neighborhood of each point on the x-abcissa. The number of points in each such neighborhood is usually taken to be about 30% of the total number of observations. Weights diminish by a cubic function, so points very near to the selected point get by far the most weight. The individual fitted regression line determines only the y-value for the selected abcissal value. In effect, the technique behaves much like a moving average, but has various advantages.

In principle, one might extend the test of linearity to the lowess technique, but the fitting procedure is such that there is no way to calculate an effective number of degrees of freedom. We can, however, utilize the bootstrapping technique to see whether the two sums of squares (about the loglinear regression and about the lowess curve) seem likely to differ significantly. Two approaches to bootstrapping regression data have been described by Efron and Tibishirani (1993). One approach depends on taking a random sample, with replacement, of the observed data point-pairs and recalculating the regression line. This method may not work well with the data considered here unless there are a substantial number of points. With fewer data points, repeated selection of the same data point occurs often enough so that the lowess process fails. In the second approach, one fits a regression line and calculates deviations from the line. Random samples of the deviations are then taken, again with replacement, and these deviations are then added to the values calculated from the original regression line at each of the observed x-values. Regression lines are then calculated from the new data sets. This method works satisfactorily with a limited number of data points, and thus might be used to compare a loglinear fit with one obtained by the lowess method.

9.7 Confidence limits from bootstrapping

The usual linear regression model assumes a normal distribution of deviations about the regression line as a basis for calculating confidence limits and tests of significance. The logarithmic transformation does seem to usually result in a symmetric distribution of deviations from regression, but it may be worthwhile to use bootstrapping to obtain confidence limits that do not depend on the assumption of normality. As noted in Section 9.6, Efron and Tibishirani (1993) describe two approaches to bootstrapping in regression. One, which they call "parametric" bootstrapping, depends on fitting the usual regression model to calculate individual deviations from regression, and then taking random samples of these deviations with replacement (same sample size as the original data) and adding this set of deviations to the y-values predicted by the original regression equation to generate a new set of "bootstrap data". A regression line is fit to this data and the parameters recorded. Doing this, say, 1,000 times provides the basis for calculating confidence limits on the parameters. For 95% limits we use the 25th and 975th ordered value of a parameter. The original set of x-values is used throughout. This approach thus makes it feasible to work with small data sets -- say 5 to 10 x,y pairs. A disadvantage is that the method assumes that the linear model is exactly true.

An alternative ("nonparametric" regression bootstrapping) takes n random samples with replacement of the x,y pairs and computes new regressions. The disadvantage is that, in small data sets, the sample

observations may pile up on a few x-values, yielding strange results. The advantage is that the method does not assume the linear model holds exactly. The two methods were tested on a number of data sets from increasing and decreasing populations, and compared with confidence limits obtained from the usual regression equation. The nonparametric approach gave results (Fig. 9.10) that agree very well with the usual 95% regression limits on slope of the loglinear regression line. Most of the samples had $n \geq 10$, but two feral horse populations had only 5 and 8 observations, respectively (Table 9.5). The nonparametric interval for brown bears is somewhat smaller than that from the standard regression calculations, and this may reflect the fact that the brown bear population may not have been changing over the years studied (Fig. 9.8).

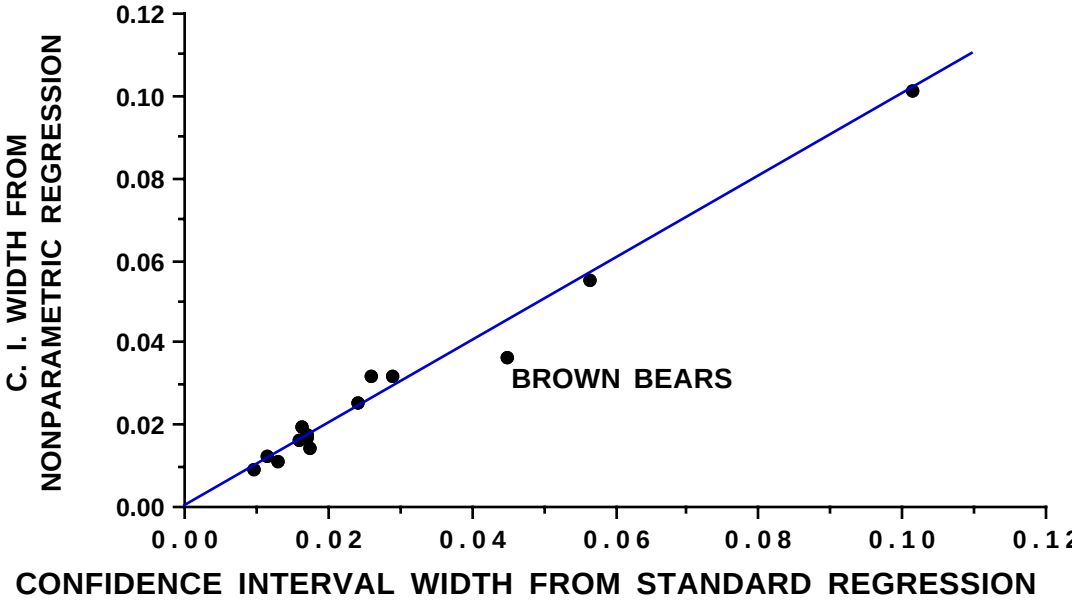

Fig. 9.10 Comparison of 95% confidence intervals on the regression slope (b) calculated by ordinary regression methods to those obtained from nonparametric regression.

Using parametric regression bootstrapping with the same data sets gives confidence limits appreciably smaller than those from standard regression (Fig. 9.11). The question of which set of limits to use needs further consideration, but the good agreement of the nonparametric and standard regression limits might lead one to prefer either over the parametric regression, particularly in that these two approaches are more conservative (wider limits) than the parametric approach.

When sample sizes are, say, 10-15 or larger, a worthwhile approach is to compute confidence limits both by nonparametric regression bootstrapping, and by the usual approach (readily available in EXCEL). If they agree, then there should be little reason for concern. Both of the bootstrap methods support the notion that regression estimates of population rate of change are unbiased. Efron and Tibishirani (1993) indicate that close agreement between the mean of bootstrapped data and the value obtained from the original approach indicates an unbiased estimate. Table 9.5 shows that the 3 methods give essentially the same values for rate of increase.

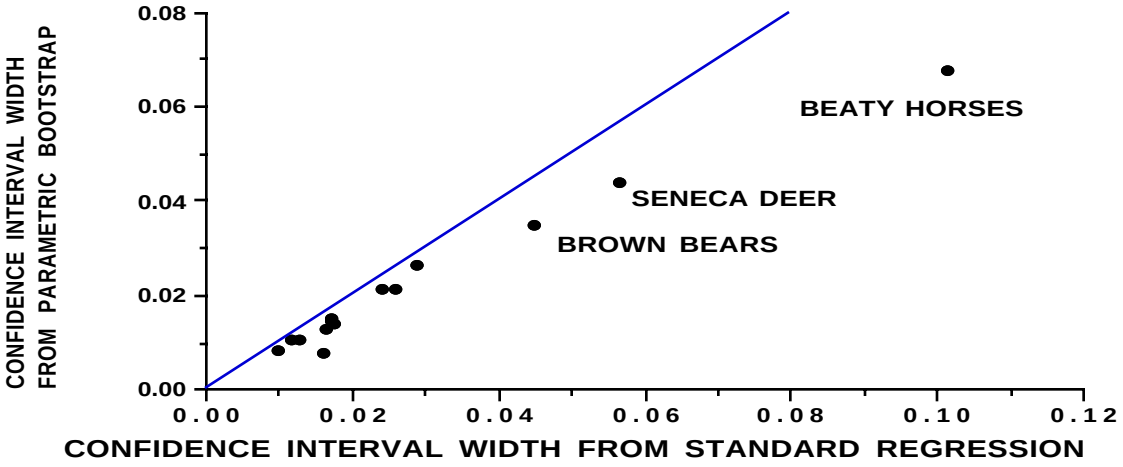

Fig. 9.11 Comparison of 95% confidence intervals on the regression slope (b) calculated by ordinary regression methods to those obtained from parametric regression.

Table 9.5 Slope estimates from log-linear regression compared to those from bootstrap samples of parametric and nonparametric regressions. The close agreement indicates unbiased estimation (Efron and Tibishirani 1993).

SPECIES	SLOPE	PARA MEAN	NPARA MEAN
OTTERS	0.0521	0.0523	0.0524
GRAY WHALES	0.0257	0.0259	0.0259
CRYSTAL R	0.0968	0.0971	0.0975
BLUE SPGS	0.0793	0.0791	0.0793
SENECA DEER	0.4115	0.4151	0.4138
MUSKOX	0.1463	0.1461	0.1465
GRAY SEALS	0.0741	0.0742	0.0743
BISON	0.2068	0.2068	0.2064
FUR SEALS	0.0824	0.0825	0.0833
BEATY HORSES	0.2447	0.2438	0.2438
LISIANSKI	-0.0260	-0.0260	-0.0257
PRYOR HORSES	0.1854	0.1855	0.1854
BROWN BEARS	0.0167	0.0165	0.0159
FUR SEAL DEC	-0.0263	-0.0264	-0.0261

9.8 Alternative estimates of rate of population change

In Chapter 12 two very simple models (Section 12.4) are proposed for evaluating populations from which known numbers of individuals are removed at various times. The log-linear approach of the previous sections (eq. 9.2) is not useful in such circumstances, so it is worthwhile to consider alternative ways to establish rates of change. These make use of ratios of successive observations. We first consider using the methods for data without removals, and thus can compare them with the regression approach. The basic idea comes from the simple relationship (eq. 9.1):

$$N_t = \lambda N_{t-1}$$

where λ is the multiplier needed to project a population at time t-1 to time t (we assume an interval of one year, with observations taken at the same time each year). If we want to estimate λ from a sequence of years, we can consider 3

ratio estimates (Eberhardt 1987) where x_i represents population size in one year and y_i the size in the next year:

The mean of individual ratios $$\hat{\lambda} = \frac{\Sigma(y_i/x_i)}{n} \quad (9.7)$$

Ratio of sums $$\hat{\lambda} = \frac{\Sigma y_i}{\Sigma x_i} \quad (9.8)$$

Regression through the origin $$\hat{\lambda} = \frac{\Sigma y_i x_i}{\Sigma x_i^2} \quad (9.9)$$

Because the individual observations other than the first and last appear twice in a sequence of years, one can get spurious results in statistical analysis of such data (Eberhardt 1970). We thus resort to bootstrapping here. Using the data previously used to evaluate the regression approach (Table 9.5), 1,000 bootstraps were used with the estimates of eqs. (9.7) through (9.9). Bootstrap bias calculations were made, along with comparisons with the rate of change from the regression method (using $\lambda = e^b$, where b is slope of the loglinear regression). All 4 sets of estimates were highly correlated (Table 9.6).

Table 9.6. Correlations between four methods of estimating rate of increase.

	LOGLINEAR	MEAN	SUMS	REGR
LOGLINEAR	1			
MEAN	0.992	1		
SUMS	0.983	0.986	1	
REGR	0.957	0.961	0.990	1

The mean of the individual ratios had the smallest relative bias. and deviated the least from the loglinear regression estimates. Relative bias was calculated as:

$$\text{Bias} = \frac{\lambda_{orig} - \lambda_{boot}}{\lambda_{orig}} \quad (9.10)$$

where λ_{boot} is the mean of 1,000 bootstrap estimates and λ_{orig} is the estimate from log-linear regression. The average of absolute values for relative bias was 0.020 for the method of means [eq.(9.7)], 0.107 for method of sums [eq.(9.8)], and 0.130 for the regression method [eq.(9.9)].

Confidence intervals (95%) were also calculated from bootstrap results, but were much wider than the regression estimates and poorly correlated with those estimates (Fig. 9.12). Excluding the Beaty horse data (n = 8), the ratio of confidence intervals (Fig. 9.12) was about 3.5 to 1, i.e., confidence intervals from the mean of individual ratios [eq.(9.7)} were nearly 4 times wider than those from log-linear regressions. This raises an issue that needs further attention. Why should the bootstrapping confidence intervals be so much wider than those from linear regression? The answer appears to lie in the fact that bootstrapping is based on random samples with replacement of *individual* observations. In this case, we have to deal with successive pairs of observations in order to get the ratios, so the pairs were bootstrapped, and this gives unsatisfactory results, as shown in Fig. 9.12.

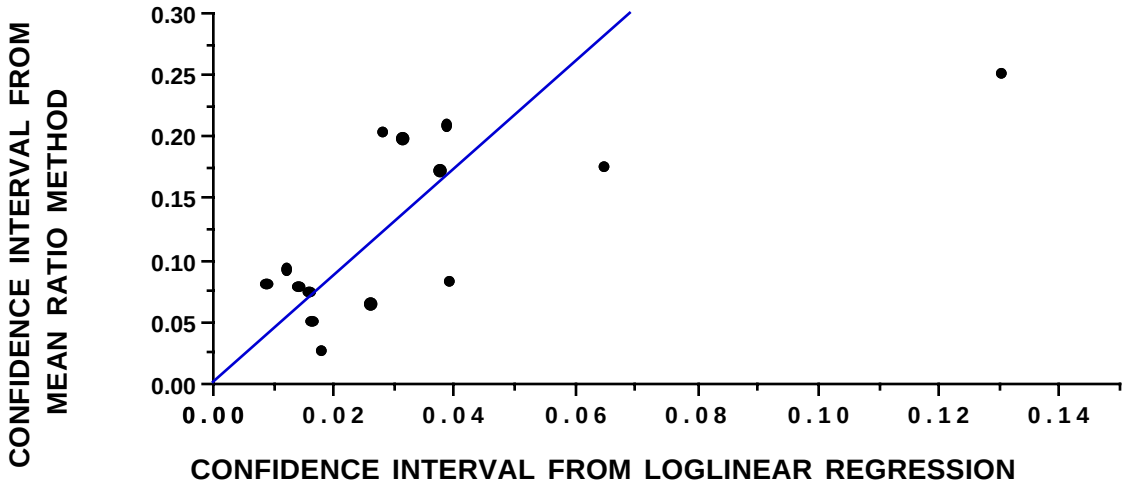

Fig. 9.12. Relationship between confidence interval widths (95%) calculated by log-linear regression and those from a ratio method [Eq. (9.7)}.

Example 9.5 A census using ratio estimation

An example of ratio estimation in an actual census is provided by studies of the abundance of the sea otter (*Enhydra lutris*) along the California coast. Sea otters then occupied about 150 miles of the California coastline, and are almost always found swimming or resting (in kelp beds, usually) just off the shoreline. A number of aerial censuses were conducted by employees of the California Department of Fish and Game. Since only a fraction of the otters present are observed in aerial counts, an ingenious use of auxiliary counts (devised by D.J.Miller) was used to correct the aerial counts by what amounts to ratio estimation. On a sample of shoreline areas, observers on vantage points (well above sea level, where possible) made counts on well-defined areas. These were then used to correct aerial counts of the same areas.

In later development of the method, aerial photos were made in advance of the census to map kelp beds. Using the maps, shoreline ("ground truth") counters located animals in a well-defined and readily visible area, and made a record of those animals present in the sample areas at the time the aerial observers passed the counting site. The aerial observers also plotted all otters seen on identical maps. As soon as a day's flight was completed, the aerial and ground observers went over the maps together to establish which animals were seen by both air and ground observers, and which animals were not seen from the air. Aerial counts were made over the entire coastline occupied by otters, so the data can be used in a ratio estimate (Sec. 4.12).

The estimated total otter population is:

$$\hat{Y}_R = \frac{\Sigma y_i}{\Sigma x_i} X_T$$

where Y_R is the estimate of total otter population, and X_T is the total aerial count. The ground counts (y_i) are summed over the sample (n) of ground-truth areas, and divided by the sum of the aerial counts (x_i) on the same areas. If it is assumed that the

ground counts are made without error (an assumption that needs further checking), then the data conform to the standard conditions for ratio estimation with the exception that X_T (the total aerial count) is subject to sampling error. Calculations are otherwise straightforward, and as given in Ch.4. Table 9.5 gives data from the 1974 survey. A total of 897 otters were counted from the air, so we estimate the total population from the equation given above as:

$$\hat{Y}_R = \frac{332}{202}\ 897 = 1474$$

and find the squared coefficient of variation to be:

$$CV^2(\hat{Y}_R) = \frac{(1-f)}{n}\ [c_{yy} + c_{xx} - 2c_{yx}] = \frac{1}{31}\ [1.966 + 1.216 - 2(1.4229)] = 0.0109$$

where we take f = 0, since an accurate value is not available. However, an appreciable segment of the coast was included in the ground counts, so that use of f would be appropriate here, if it could be calculated.

Confidence limits are readily calculated by computing a standard error as $(.0109)^{1/2}1474 = 159.89$, and using $Z_{05} = 1.96$ giving:

$$\hat{Y}_R = \pm Z_{05}\ s(Y_T) = 1474 \pm 1.96(159.89) \quad \text{or,}$$

$$1172 < Y_R < 1776 \text{ sea otters.}$$

Aerial survey (x_i) and ground counts (y_i) of sea otters along the California coast.

June 25		June 26		June 27	
x_i	y_i	x_i	y_i	x_i	y_i
2	4	1	6	1	1
1	1	9	10	6	14
0	4	0	19	5	7
0	6	13	19	1	6
37	50	12	8	4	5
6	9	5	8	2	8
9	10	6	6	2	1
3	3	5	9	37	47
1	1	0	5		
8	13	1	1		
0	8	11	14		
		14	29		
67	109	77	134	58	89

9.9 Criteria for regression fits

Testing for curvilinearity in regression lines was discussed in some earlier sections (Sec. 9.3, Sec. 9.5). When one fits a regression line like those summarized in Fig. 9.4, it is important to have other ways to evaluate the fit. Regression mean squares provide a useful measure of variability about the line, and can be supplemented by the widely-used R^2 criterion. For a simple regression line with one independent variable, the square root of this quantity gives the well-known correlation coefficient, and R^2 is perhaps better known when used in the multiple regression analyses that will be

$$R^2 = \frac{\Sigma(\hat{y}_i - \bar{y})^2}{\Sigma(y_i - \bar{y})^2} = 1 - \frac{\Sigma(y_i - \hat{y}_i)^2}{\Sigma(y_i - \bar{y})^2} \qquad (9.11)$$

described in a later section. However, the simple linear regressions serve to illustrate the interplay between regression mean square and R^2. The R^2 criterion is:

where y_i is the independent variable, $\bar{y}$ is the mean of the y_i and $\hat{y}$ is the "predicted" value from the regression equation. Often, R^2 is described as the proportion of the variance in the independent variable "accounted for" by the fitted regression line. The two expressions in eq.(9.11) are connected by the basic identity in the analysis of variance in regression, i.e.,

$$\Sigma(y_i - \bar{y})^2 = \Sigma(\hat{y}_i - \bar{y})^2 + \Sigma(y_i - \hat{y}_i)^2$$

Total S.S. = regression S.S. + Residual S.S.

In eq.(9.11), the left-hand form is that commonly used (Draper and Smith 1998:138). The equivalent right-hand expression was recommended by Anderson-Sprecher (1994) because it provides a convenient interpretation of R^2 written as:

$$R^2 = 1 - \frac{RSS(full)}{RSS(reduced)}$$

where RSS(full) denotes the regression sum of squares for the full model and RSS(reduced) can be interpreted as the sum of squares for the model reduced to its minimal form, i.e., the expected value of y is β_0, estimated by $\bar{y}$. One advantage of this expression is that it emphasizes that model comparisons using R^2 should be made with nested models, i.e., a series of regression equations with two or more independent variables so that the number of parameters p = 2,3,4, ... Another advantage is that

$$1 - R^2 = \frac{RSS(full)}{RSS(reduced)}$$

states the fraction of variability <u>not</u> accounted for by regression. Because at least stochastic fluctuations are always present in trend data, this expression serves as a reminder that R^2 cannot become unity.

When R^2 is large, it is evident that the regression line does a good job of predicting the counts. This does not necessarily demonstrate validity of the

index as some extraneous factor may be exerting a major influence on the counts. Usually, however, a high R^2 is reassuring. A key element in demonstrating validity of an index is an independent estimate of the trend. Such an estimate may be available from reproductive and survival data. Thus if two sources, trend index and reproductive and survival data produce much the same estimate of λ, that result is particularly reassuring.

Confidence intervals on parameters and on predictions from regression lines depend on the assumption of a normal distribution of deviations from the regression line with constant variance about regression. When population data are being considered, this may be an uncertain assumption. It is thus desirable to resort to the bootstrapping technique (Efron and Tibshirani 1993) as a check on confidence intervals generated by regression theory. For regressions based on modest numbers of data points, a "parametric regression" bootstrap is recommended. The usual regression line is fitted and deviations from that line are bootstrapped. That is, at least 1,000 random samples of size n (n = no. of points used to fit the regression line) are drawn from the set of deviations, with replacement, and are used to construct new regression data sets (by adding the sample of deviations to the estimated regression line). A new regression line is fitted to each such data set and used to estimate the parameter of interest (usually λ). Confidence intervals are then obtained by counting in $B\alpha/2$ values from either end of the generated distribution of values, where B = no. of bootstrap samples and α denotes the chosen level of significance (often $\alpha = 0.05$).

If bootstrapping is used, it furnishes an estimate of bias for the estimator of concern from the equation (Efron and Tibshirani 1993):

$$\widehat{\text{bias}}_B = \hat{\theta}^*(\bullet) - t(\hat{F}) \qquad (9.12)$$

Here, $\hat{\theta}^*(\bullet)$ is the mean of the bootstrap estimates of the parameter of interest (often λ) and $t(\hat{F})$ denotes the same parameter estimated from the original data. Results from this criterion may thus yield an indication of problems with the trend index.

How should these two criteria be used? An interpretation of R^2 was given earlier as:

$$R^2 = 1 - \frac{\text{RSS(full)}}{\text{RSS(reduced)}}$$

Here RSS denotes a regression sum of squares. The numerator pertains to the fitted model, while the denominator consists of the variance about the mean of the observations (thus "reduced" to a minimum). Usually the sample size is large enough so that the ratio amounts to comparing the variance of the fitted model to that in the data (i.e., the degrees of freedom are not different enough to matter much). Hence it is clear that R^2 and the regression mean square are closely related. The relationship can be examined by plotting (Fig. 9.13) R^2 and regression mean squares for the data of Fig. 9.4.

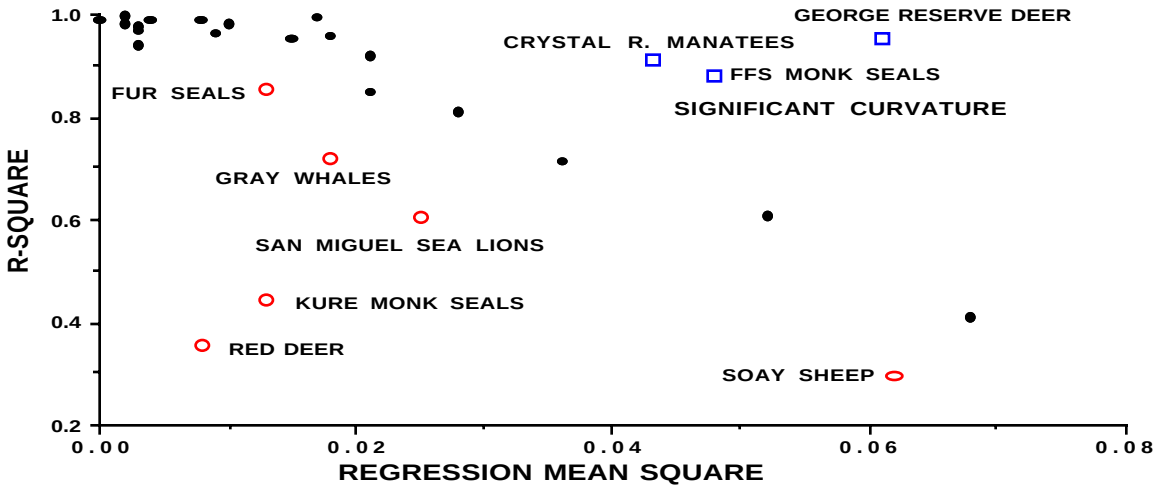

Fig. 9.13. Relationship between R^2 and regression mean square for data used in Fig. 9.4. Various sets of points described in text.

The solid points show what seems likely to be the expected relationship between R^2 and the regression mean square. The cluster at the right top contains those data sets where there is a significant curvature which inflates the regression mean square. The points represented by open circles represent cases where there appears to be a pattern in the deviations which is quite dramatic in some instances. Plotting deviations from regression, as recommended earlier (see Fig. 9.7 and Fig. 9.14), will usually make the non-randomness of the deviations evident. All of the cases shown here are based on simple log-linear regressions. Very likely a detailed investigation of the underlying circumstances might turn up significant auxiliary variables. Considering the solid points at the left of the figure, it appears that most of the variability is accounted for by the simple regression. R^2 thus appears most informative in this situation.

Evaluating the pattern in deviations from regression can be aided by the Durbin-Watson test (Draper and Smith 1998). This simple test depends on the fact that the squared difference between successive deviations will approximate the variance of the deviations if the pattern of deviations is random. The test is:

$$d = \frac{\Sigma\ (e_u - e_{u-1})^2}{\Sigma\ e_u^2} \qquad (9.13)$$

where the summation in the numerator runs from $u = 2$ to n, and that in the denominator runs from 1 to n. It can be shown that the ratio, d, has an expected value of 2 under a random pattern. Draper and Smith (1998) note that $0 \leq d \leq 4$, and give tables of significant deviations from the expected value of 2 for different sample sizes. Various statistical packages produce the Durbin-Watson test on residuals, but the test is easy to compute and thus worth calculating directly once one has the residuals from regression (readily available with the spreadsheet regression calculations). Three of the sets of data in Fig. 9.13 show significance at the 1% level from the Durbin-Watson test.

The pattern of deviations for these three species appears in Fig. 9.14, where the correlation of successive observations is quite evident.

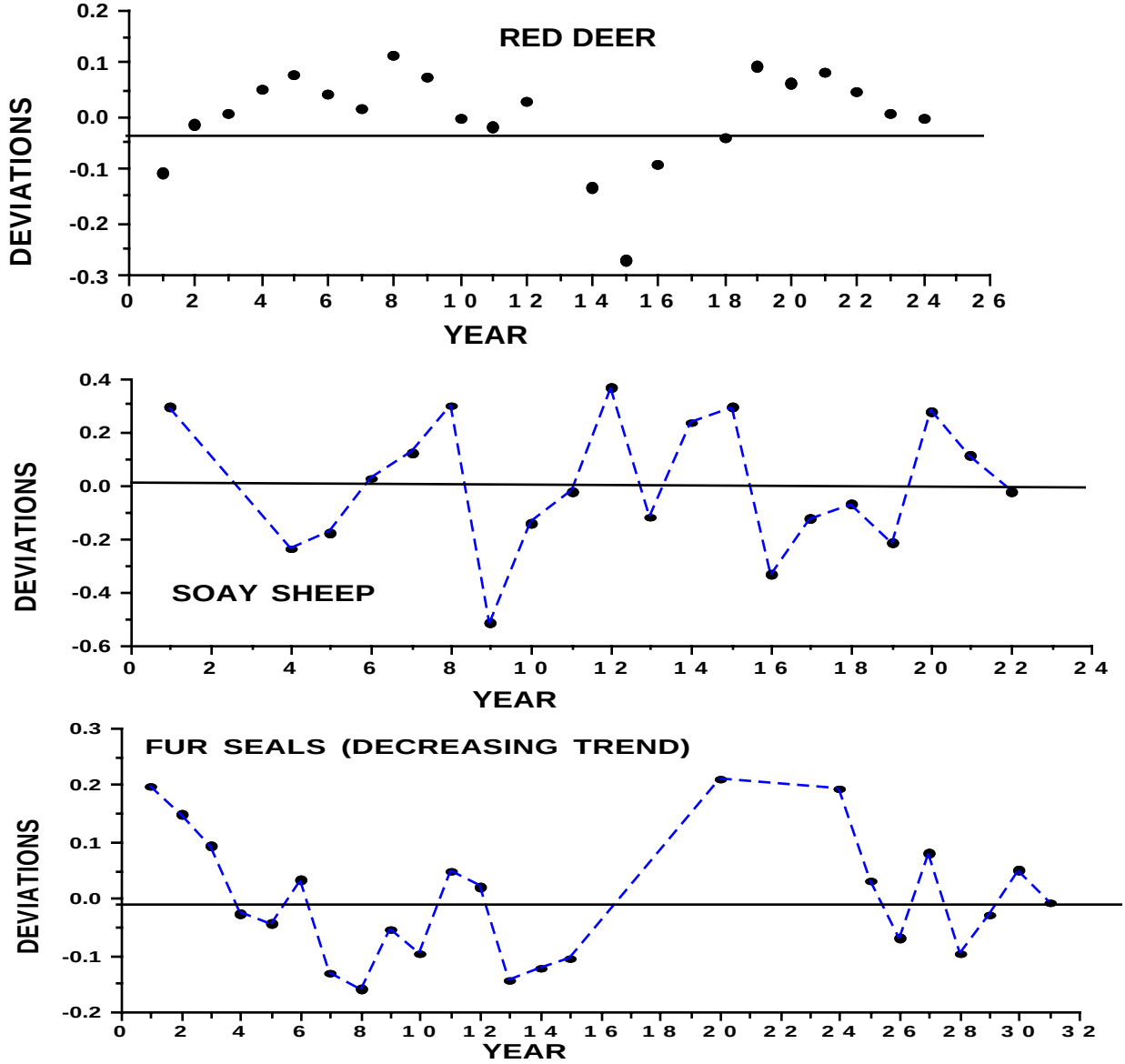

Fig. 9.14 Pattern of deviations from loglinear regression for 3 species. All 3 sets of data are significant at the 1% level with the Durbin-Watson test for serial correlation.

The results above thus suggest several steps in appraising loglinear regressions of trend data:

1) Test the data for curvilinearity.
2) Compute the regression mean square and R^2.
3) Examine the pattern of deviations from regression over time, and calculate the Durbin-Watson test..

It may be helpful to compare the results with the data of Fig. 9.13. Most of the data points there conform to the above steps, but not all. Three of the open circles (Kure monk seals, San Miguel sea lions and gray whales) do not fit

in, but plotting the residuals suggests that the data are quite erratic, and that there may be some other factors involved that need to be further investigated. Significant curvilinearity (3 points on the upper right of Fig. 9.13, Crystal River manatees, French Frigate monk seals, and George Reserve deer) also indicates a need to look further at the data, inasmuch as the curvilinearity may well indicate a significant change in trend. Table 9.7 gives the data on species shown in Fig. 9.13. The very high R^2 values shown by a number of the data sets suggest that the variability is mainly a function of stochasticity. Bison, musk oxen, Cusino deer and Pryor wild horse populations were essentially counted in their entirely, while the Seneca deer population was reconstructed from removals that were known almost completely.

Table 9.7 Data on loglinear regressions used in Fig. 9.17. Data ordered by regression mean squares. Population sizes are rough estimates in a number of cases.

SPECIES	NUMBERS	OBSNS	SLOPE	Lambda	MSreg	R-sq
GRIZZLY BEARS	400	18	0.039	1.040	0.068	.409
SOAY SHEEP	1000	20	0.026	1.026	0.062	.295
GEORGE RESERVE DEER	70	7	0.488	1.629	0.061	.956
PEARL & HERMES MONK SEALS	150	16	0.050	1.051	0.052	.606
FFS MONK SEALS	600	22	0.062	1.064	0.048	.882
CRYSTAL RIVER MANATEES	150	21	0.097	1.102	0.043	.912
LISIANSKI MONK SEALS	300	35	-0.029	0.971	0.036	.714
LAYSAN MONK SEALS	300	34	-0.034	0.967	0.028	.812
SAN MIGUEL SEA LIONS	8000	15	0.064	1.066	0.025	.603
WOOD BUFFALO N.P. BISON	5000	19	-0.052	0.949	0.021	.849
BLUE SPRINGS MANATEES	40	19	0.079	1.082	0.021	.92
BEATY HORSES	400	8	0.245	1.277	0.018	.959
GRAY WHALES	15000	18	0.026	1.026	0.018	.718
SENECA DEER	400	11	0.411	1.508	0.017	.992
ALE ELK	70	11	0.165	1.179	0.015	.956
FUR SEALS (DECREASING)	180000	24	-0.026	0.974	0.013	.851
KURE MONK SEALS	100	12	0.026	1.027	0.013	.443
LONGHORN CATTLE	20	7	0.302	1.353	0.010	.981
YELLOWSTONE ELK	7000	8	0.191	1.210	0.009	.964
RED DEER	1600	22	-0.010	0.990	0.008	.354
BISON	160	14	0.207	1.230	0.008	.99
MUSKOX	120	15	0.146	1.157	0.004	.991
SERENGETI BUFFALO	50000	11	0.064	1.066	0.003	.969
CALIFORNIA SEA OTTERS	1600	13	0.052	1.054	0.003	.94
GRAY SEALS	1200	16	0.069	1.072	0.003	.975
CUSINO DEER	80	5	0.370	1.448	0.002	.995
FUR SEALS (INCREASING)	130000	12	0.082	1.085	0.002	.984
PRYOR HORSES	120	5	0.185	1.204	0.000	.99

9.10 Using auxiliary variables with trend data

In some cases, the use of auxiliary variables may serve to reduce the regression mean square, i.e., we add independent variables other than time and use a multiple regression equation. One such model was used to study trends in manatee numbers by Garrott et al. (1994,1995). They represented the

expected number of manatees counted at any given site under average conditions at time t as M(t), and assumed M(t) remains constant over the annual counting period (given average conditions) and denoted it as M_i for the ith year. M_i is assumed to be proportional to the true population level. C[t,X(t)] then represents the expected number of manatees counted at time t, given conditions X(t), where X(t) is a vector of covariate values prevailing at time t, leading to the model:

$$C[t,X(t)] = M_i\, R[t,X(t)] \tag{9.14}$$

where the function R( ) is a rate function that takes the value unity when conditions are average. Under good counting conditions R( ) > 1 and under poor conditions R( ) < 1. If the rate function is assumed to be of the form

$$R[t,X(t)] = \exp[x(t)'\beta]$$

where x(t) is the vector of covariates and β is a vector of regression parameters, then taking logarithms (base e) gives:

$$\ln C[t,X(t)] = \ln M_i + x(t)'\beta$$

One can then use multiple regression to study the effect of various covariates (auxiliary variables). The general formulation is widely used in survival studies as "Cox's proportional hazards model" (Cox 1972) and has also been used to take auxiliary variables into account in population estimation procedures. The main interest in eq.(9.14) is for studying population trend. In order to do so, M_i needs to be expressed as a function of time, usually as $M_i = N_0\lambda^t$ so that the final equation becomes:

$$\ln C(t) = \ln N_0 + t \ln \lambda + x(t)'\beta \tag{9.15}$$

Compare this with eq.(9.1) and (9.2).

The underlying model is thus assumed to have the form:

$$y = \beta_o + \beta_1 x_1 + \beta_2 x_2 + \beta_3 x_3 + \ldots + \beta_{p-1} x_{p-1} \tag{9.16}$$

where $y = \ln C(t)$, $x_1 = t$, $\beta_o = \ln N_o$, $\beta_1 = \ln \lambda$, x_2 .... x_{p-1} are the auxiliary variables, and there are p parameters to fit with multiple regression.

The use of regression mean square, R^2 and a bias criterion were discussed in Sec. 9.9. Two further criteria have been used for evaluation of models, Mallow's Cp for multiple regression models (Draper and Smith 1981) and Akaike's Information Criterion (AIC) for models where likelihood ratio tests are appropriate (Lebreton et al. 1992, Burnham and Anderson 1996). The C_p statistic is calculated as (Draper and Smith 1998:332):

$$Cp = RSSp/s^2 - (n-2p) \tag{9.17}$$

where RSS_p is the residual sum of squares in a multiple regression model based on p parameters, n is the number of observations and s^2 is the residual mean square from the equation with the largest number of parameters in the set of equations evaluated. The method thus depends on having a range of auxiliary variables available for study and essentially assumes that this set of variables includes those involved in the "true" underlying regression model. Draper and

Smith (1998:331) suggested plotting residual mean squares against the number of parameters (p) for a sequence of regression models as a way to estimate an asymptotic value that may approach the "true" value σ^2, which is then used as s^2 in eq.(9.17). They noted that large samples and a sizable number of candidate variables should be available for this approach to be valid.

Burnham and Anderson (1996) gave the Akaike (AIC) criterion as: AIC = -2(log-likelihood) + 2p where p is again the number of parameters and the log-likelihood ratio is calculated from the maximum likelihood estimates of parameters in two candidate models. In theory, the method requires that a "global" set of models be identified and that this set contains the "true" model according to Burnham and Anderson (1996). A series of applications of AIC for survival analysis was provided by Lebreton et al.(1992) and Burnham and Anderson (1996) provided a further example. Because regression models are used here for analysis of indices, the Cp statistic seems useful. Burnham and Anderson (1996) discuss the analogous features of AIC and Cp. Much more detail is available in the book by Burnham and Anderson (1998) where it is claimed that the set of models considered does not need to include the "true" model.

Example 9.6 Trend indices with auxiliary variables.

Two examples of trend indices of the use of auxiliary data illustrate the approach of the previous section. One uses the data on manatees studied by Garrott et al. (1994, 1995). The other considers data on the Yellowstone grizzly bear population. Background data for both species appear in the Case Histories.

Garrott et a. (1994, 1995) evaluated a sizable number of potential auxiliary variables, but it appears that year and DD10 (cumulative heating days summed for 10 days previous to the aerial counts of manatees in warm-water refugia) may serve as well as larger sets of temperature variables (Eberhardt, Garrott and Becker1999). The model assumed for the study was the "proportional hazards" model of eq.(9.15), fitted by multiple regression [eq.(9.16)]. R^2 was about 0.60 for several versions of the overall multiple regression analyses. A difficulty with the results is that the estimated rate of growth of the manatee population exceeded that believed likely (Eberhardt, Garrott and Becker 1999), and estimated from reproductive and survival data (Eberhardt and O'Shea 1995). An alternative approach used regression of repeated counts within years on DD10 for a covariance adjustment (Snedecor and Cochran 1967), as detailed by (Eberhardt, Garrott and Becker 1999).

The alternative approach suggested that the population remained relatively constant over recent years, in accord with the conclusion of Eberhardt and O.Shea (1995). Plotting deviations from the multiple regression fit (Fig. 9.15) suggests that some factor not accounted for in the model may have influenced the trend. As a check on the use of ordinary multiple regression, a nonparametric bootstrapping study was conducted. There were 103 data points in the manatee data set. These were randomly sampled by taking n = 103 random samples with replacement and calculating a multiple regression on each such sample. Doing this 2,000 times yielded estimates and percentile confidence limits essentially the

same as those given by ordinary multiple regression. The usual regression program gave an estimate of 0.0958 for the regression slope with a 95% confidence interval of 0.074 to 0.118 while bootstrapping produced a mean estimate of 0.0953 with 95% confidence interval of 0.076 to 0.115. The bias estimate of eq.(9.12) is thus $\widehat{bias}_B$ = 0.953- 0.0958 = - 0.0005, so there is no evidence from the bootstrapping about problems with the regression approach.

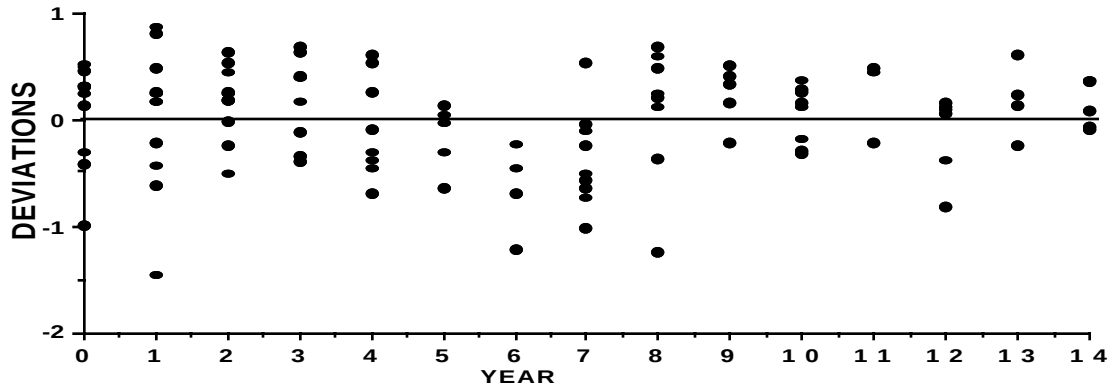

Fig. 9.15 Deviations from a multiple regression model fitted to manatee data.

The Yellowstone grizzly bear data yielded higher values of R^2 and used 3 auxiliary variables: year, frequency of sighting of given family groups, and April snow depths in the previous year. The index variable was counts of "distinct families", i.e., of females with cubs-of-the-year seen in the summer (Knight, Blanchard and Eberhardt 1995). The model thus was:

$$\ln(\text{count})=\beta_0+\beta_1(\text{year})+\beta_2(\text{frequency})+\beta_3(\text{snowpack}) \qquad (9.18)$$

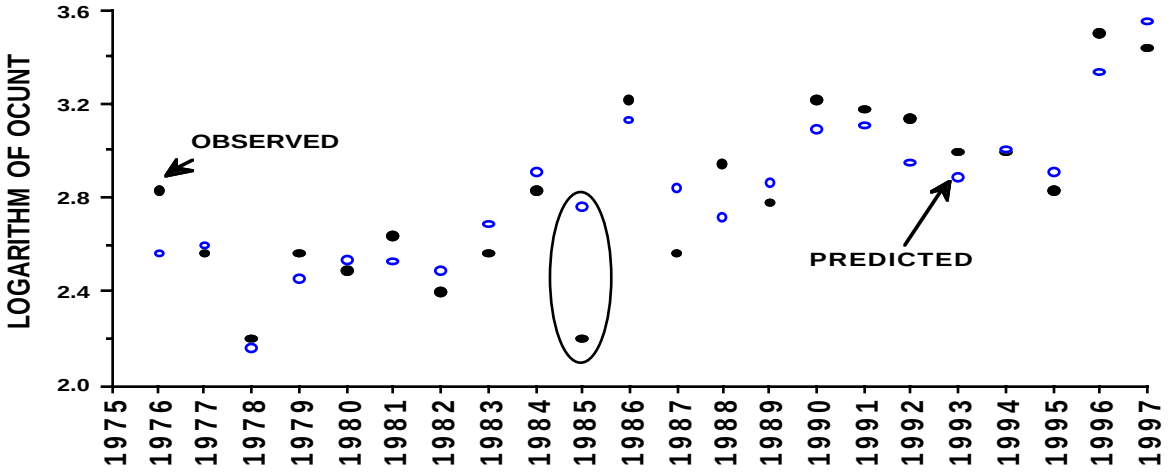

Fig. 9.16. Observed (solid points) and predicted values (open circles) of logarithms of counts of "distinct families" of grizzly bears in Yellowstone National Park and environs. Aberrent 1985 value is circled.

This model yielded an R^2 of 0.75 using data from 1976 to 1997. Comparison of values calculated from the regression equation and observed counts (Fig. 9.16) suggests that the observed value in 1985 was somehow aberrant. Without this value, $R^2 = 0.85$. A variety of additional variables were examined in multiple regressions, including squared terms for year and frequency of sighting as used in the manatee studies of Garrott et al. (1994, 1995). None of these additional variables appeared to provide useful fits. Mallow's C_p [eq.(9.17)] was calculated (Table 9.8) by plotting s^2 against number of variables included in the regression as recommended by Draper and Smith (1998:331). This suggested a value for s^2 (0.03) from the trend of calculated regression mean squares. Table 9.8 also includes $1 - R^2$ as a measure of the proportion of variance not accounted for by the regression lines. The three measures show essentially the same trend with the number of parameters estimated (p). Another reassuring aspect of the revised index is that there is now little evidence of curvilinearity in the residuals (Fig.9.17) in contrast with an earlier index calculation using only year and frequency of capture (Eberhardt, Garrott and Becker 1999). There were no significant correlations between the independent variables.

Table 9.8. Variation in three measures of regression model adequacy with increasing number of parameters (p) included in the model. Data for a Yellowstone grizzly bear trend index fitted to models of the general form of eq.(8).

p	s^2	$1 - R^2$	C_p
2	0.053	0.36	15.6
3	0.039	0.25	7.6
4	0.039	0.23	7.2

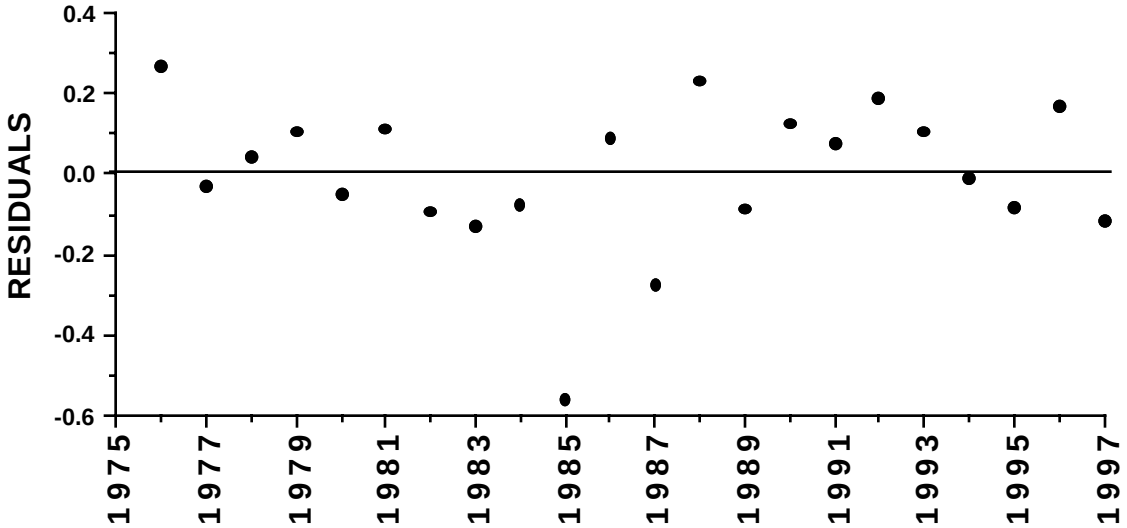

Fig. 9.17. Deviations from regression for the grizzly bear data of Fig. 9.16.

Bootstrapping was used to check the multiple regression calculations, with essentially the same results as for the manatee data. The year coefficient was 0.0287 with a 95% confidence

interval of 0.017 to 0.040, while the mean of the bootstrap calculations (2,000 replications) was 0.0291, with 95% confidence interval of 0.018 to 0.040. The bias estimate of eq.(9.12) is thus 0.0291- 0.0287 = 0.0004, so there is again no evidence from bootstrapping of problems with the regression approach. As with the manatee data, the bootstrap frequency distribution was symmetric about the estimate.

Example 9.7 An alternative approach to index models.

A different prospect for assessing trend data can be illustrated by using data on wolves and moose from Isle Royale (Peterson 1995), and a difference equation model used by Eberhardt (1998). The model is:

$$V_t = [1 + r_1]\ V_{t-1} - [\frac{r_1}{K^z}]V_{t-1}^{z+1} - cH_{t-1} \qquad (9.19)$$

Where V_t denotes ungulate prey abundance at time t, and H_{t-1} denotes predator abundance the previous year, K is the asymptotic population level of prey, z is a constant for the generalized logistic equation (Eberhardt 1987), r_1 is the maximum rate of increase of prey, and c is the predation rate (prey taken per wolf per year). The above model can readily be fitted by multiple regression, giving the results of Fig. 9.18. DelGiudice et al. (1997) indicated that the moose population on Isle Royale was importantly affected by an epizootic of the winter tick (Dermacentor albipictus) in 1989, so only the data series through 1988 is used here. R^2 for the regression fit is 0.91, suggesting a good fit.

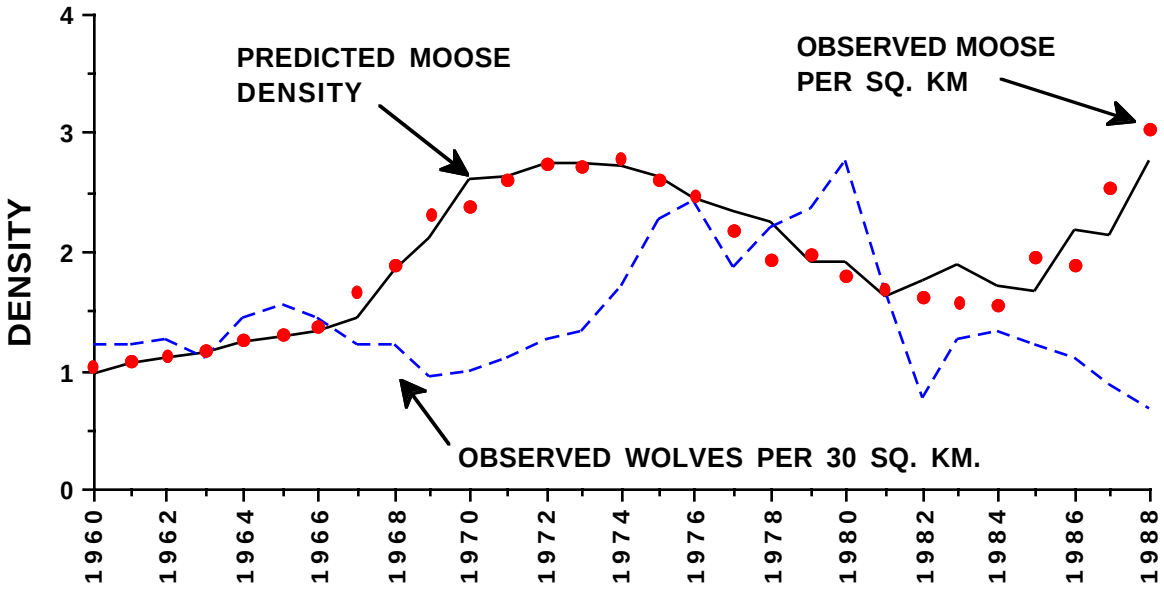

Fig. 9.18. Fit (solid line) of eq. (9.19) to observed data on moose (solid points) and wolves (broken line) on Isle Royale.

Bootstrapping was again applied to try to check validity of the approach. However, the structure of the model forces use of "parametric" bootstrapping, in which deviations from the fitted model are randomly sampled with replacement, attached to the model fitted to the original data, and refitted. This was done 2,000 times. Using nonparametric bootstrapping here poses problems, because the model is fit to observations taken sequentially.
Results of the bootstrapping study appear to support the model, giving mean values for the 3 coefficients close to those from the original fitting. The original fit estimated $\lambda = 1 + r_1$ as 1.309, and bootstrapping gave 1.314. The second coefficient in the model was -0.00126 from the original data while bootstrapping yielded an average of -0.0013. The third coefficient (c) was -7.626 in the original fit, while bootstrap data averaged -7.651. Frequency distributions of the bootstrapped data gave wider 95% confidence intervals than might be desired. Those for λ were about 1.19 to 1.44, and the interval for the moose kill per wolf (c) was -3.78 to -11.80. The interval on R^2 was somewhat more satisfactory, being 0.88 - 0.96. The Durbin-Watson test on residuals was 2.01, indicating virtually no deviation from randomness. A plot of the deviations (Fig. 9.19) does, however, emphasize the variability evident in the fit to the later years evident in Fig. 9.18.

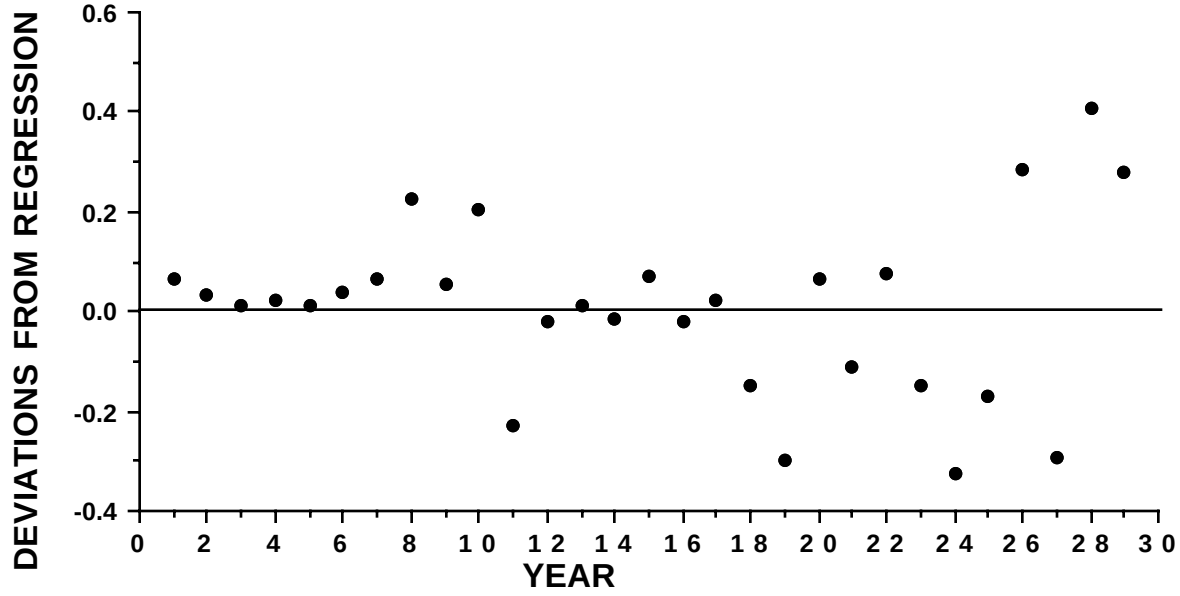

Fig. 9.19 Residuals from the multiple regression fit of eq.(9.19) to Isle Royale moose and wolf data.

One might thus be inclined to suppose that the model of eq. (9.19) may give a useful representation of the data. Unfortunately, the bootstrapping exercise may be of uncertain utility here, inasmuch as Efron and Tibshirani (1993) point out that "parametric" bootstrapping results depend on the assumption that the underlying model is correct, and this may not be true here. One problem is that the model may induce correlations, by virtue of the fact that all but one of the observations appears both in the dependent variable (V_t), and in the first independent variable (V_{t-1}).

The same kind of problem exists in other analyses of the data, but arises in a different way. Mech et al. (1987) used linear regression to relate snow accumulation to moose abundance on Isle Royale. Their results were critiqued by Messier (1991), followed by a response by McRoberts et al. (1995). The data were again studied by Post and Stenseth (1998). The difficulty is that these analyses depend on the ratio of successive population sizes (see, for example, eq.(1) and (2) of Post and Stenseth, 1998). Using such a ratio can induce correlations, as was pointed out by Watt (1964,1968) and further illustrated by Eberhardt (1970). Consequently, analyses should somehow use only the current observations as the dependent variable, and should not include the population index as an independent variable in the regressions. Further study is thus needed to determine just how to proceed. Mech et al. (1987) show various correlations between reproductive parameters and snow depth, so the question is not one of whether winter conditions have an effect, but is rather one of the impact on population trend.

9.11 Catch-effort methods

The catch-effort methods have been developed as a way to use information gained in the course of exploiting a population. By far the main use has been in connection with commercial fisheries. The methods have had very little use in appraising either sports fishing or hunting, although presumably suitable data has been increasingly available from surveys of hunting and fishing. Very much the same techniques are also encountered under the title of "removal methods", mostly in situations where animals are killed by traps, or removed from the study area for some reason.

Use of catch-effort methods in studies of exploited populations depends on the prospect for obtaining large volumes of data in return for a relatively small investment of research or management funds. Complications are, however, introduced by the nature of commercial operations. Times and places of sampling are largely not controlled by the investigators (which has led to increasing use of research vessels that can be operated in specified patterns). The exploitation is usually continuous within the seasons set by regulations, but the amount of effort may vary substantially in time, and may also be correlated with past success and local population densities. These and other difficulties have led research workers to incorporate tagging programs with catch-effort studies. Such tagging programs are almost always of the "single-recovery" type inasmuch as there is no prospect of releasing tagged individuals caught in large scale commercial operations.

In many fisheries situations, there may be more emphasis on estimating rates of exploitation than on population size. For the most part, exploitation rate is measured by estimating survival rates, and this almost always brings in the complications involved in separating fishing and non-fishing mortality rates.

A serious problem in using catch-effort methods is that the "catchability" may change as exploitation continues. One obvious prospect is that the vulnerability of several age-classes to exploitation may differ. As the more vulnerable age-groups are removed, catchability will appear to decrease.

Normally such an effect can be studied by examining data pertaining to different age groups separately.

In sports hunting there may be two facets of changes in catchability. One is the higher vulnerability of younger animals, which leads to an apparent decrease in catchability as the season progresses. The other is that heavy hunting effort usually occurs early in the season. In some situations this may mean that hunters tend to interfere with each other and thus reduce the effective catchability early in the season. An additional factor may be that inexperienced hunters tend to give up after a few days; in combination these two factors increase catchability as the season goes on. This plus decreases in vulnerability could make for marked changes in time. However, there are also circumstances, such as deer-hunting in heavy cover, where the higher levels of effort may actually be more efficient in finding and harvesting the available animals.

Apparent changes in catchability may also be due to other circumstances. With relatively short seasons and high effort, sports hunting studies may be conducted as though the population were "closed" to other losses with the exception of "crippling" loss whereby animals are killed but not recovered. For the most part, such losses tend to be proportional to the recovered kill so that the recorded catch per unit of effort is less than actual, and the population size is underestimated. In any case, the effects of changes in catchability and other uncertainties have largely limited application of catch-effort methods to commercial fisheries. The approach should, however, be considered as a potential index method in other situations, so that the main features are discussed here.

9.12 Models for catch-effort data

The basic model for catch-effort studies is one developed by Leslie and Davis (1939) for animal trapping studies and by DeLury (1947, 1951) for fisheries work, and is thus sometimes called the Leslie-DeLury model. Work by Moran (1951), Zippin (1956), Ricker (1958), Chapman (1954), and Hayne (1949) has led to their names also being attached to various versions of the equations.

The population is assumed to be closed to all losses other than the source under study, and to any form of recruitment. One simple and useful way of approaching the method is to visualize random sweeps of a net through a fixed unit of volume in some large region containing N animals. If the sweeps are of fixed size and are randomly executed, then the prospect that an individual animal is caught in one sweep can be regarded as a binomial-type situation, with the probability of capture (P) depending on the fraction of the total volume swept by the net. It is assumed, of course, that the animals are unable to escape the net and that the sweeps all constitute the same fraction of the total volume. If the sweeps are randomly located, there is no need to assume that the animals are distributed in any particular fashion -- the probability that any given individual is caught does not depend on position of the others if the sweep is done "at random". Such an argument does, of course, have to encompass the very unlikely prospect that all of the animals present could be accomodated in one sweep.

Assuming that the sweeping is done at random, we can determine the probability that a given individual is caught on the i^{th} trial very simply as $P(1-P)^{i-1}$, that is, the animal escapes i-1 sweeps, each with probability (1-P) and is caught (with probability P) on the ith trial. Again this pertains only to the fate of a particular individual. If, however, it is postulated that the chance that any one individual is caught does not depend on the fate of the other individuals, then the expected value of the total catch, C_i, in the i^{th} sweep can be written as:

$$E(C_i) = NP(1-P)^{i-1} \tag{9.20}$$

and the equation can be converted to a model relating number of captures in each sampling unit (sweep of the net) to the accumulated units, i.e.:

$$\log C_i = \log(NP) + (i-1)\log(1-P) \tag{9.21}$$

and this equation can be fitted to data by ordinary regression methods. This is the form in which the equation has been used to describe "removal" trapping.

DeLury used a "catchability coefficient", k, rather than P, and considered results in terms of the catch per unit of effort, rather than as catch per sweep of a single net, as above. This is a change necessitated by the continuous nature of a commercial fishery, in which there may be a large number of nets or other fishing "gear" in use simultaneously, and in which the records may be in terms of summaries for fixed lengths of time. It is thus necessary to assume that the various units of effort (which may be, for example, several vessels fishing for a week) are independent (i.e., vessels do not interfere with each other's success) and write the relationship as:

$$C_t = kN(1-k)^{E_t}$$

where E_t represents the cumulated units of effort up to the time when measurement of C_t began; that is, E_t is made as nearly analogous to (i-1) in eq.(9.6) as possible. Also, when k is small and E_t is large, the equation can accurately be approximated by:

$$C_t = kNe^{-kE_t} \tag{9.22}$$

and written in logarithmic form as:

$$\log C_t = \log(kN) - kE_t \tag{9.23}$$

which can be fitted by simple linear regression of logarithms of catch per unit of effort on cumulative effort. It should be noted that the same regression fit can be used for eq.(9.21), with the main difference being in how one interprets the regression slope. In the present equation, the slope estimates k, the catchability coefficient directly, while in eq.(9.21) it estimates log(1-P). Of course, if P is small (as it will be when large populations are involved) there may be no practical difference, since log(1-P) is susceptible to series expansion in which the main term is -P. Seber(1983:302-303) calls eq.(9.21) Ricker's method, and eq.(9.23) DeLury's regression model.

A serious theoretical limitation on eq.(9.20) is that the successive C_i are by no means independent, a fact emphasized by Moran(1951). If the capture of any individual animal is assumed to be independent of that of any other individual, then a binomial model for the capture of X out of N individuals can be used:

$$P_r[\text{capturing X individuals in one trial}] = \binom{N}{x} P^x (1-P)^{N-x}$$

and the model can be extended to cover n trials in which each of C_i individuals are caught. Estimating equations were obtained by Moran(1951) and various approximate solutions and methods for obtaining sample size and so on were obtained for these "removal" methods by Zippin (1956,1958).

Another way to approach the problem of non-independence of successive catches is to consider each such catch in terms of a "conditional" model. If the probability of catching a given individual is regarded as k (as in DeLury's development) then on the i^{th} trial the expected catch is that of the binomial expectation:

$$E(C_i) = k[N - \sum_{j=1}^{i-1} C_j]$$

where the term in brackets represents the number of individuals surviving in the population up to the time of the ith sampling. DeLury generalized this model to represent an average catch per unit of effort, C_t and used K_t to represent cumulative removals up to the time period being considered so that:

$$C_t = kN - kK_t \qquad (9.24)$$

This equation can also be fitted to data by simple linear regression methods. We thus have two elementary models, eq.(9.20) and eq.(9.22), that can be applied to data from populations that are "open" only to the removal method being used to estimate population size. Applications mostly have to be limited to rather short time periods and conditions where some other form of loss quite surely does not apply. In the great majority of cases, one has to assume losses from other causes, and often to account for various forms of recruitment to the population under exploitation. These problems, plus uncertainty as to the constancy of k, the "catchability coefficient", may require special auxiliary studies, often accomplished by tagging a number of individuals in the population being studied. Those facing such problems should consult the book by Ricker(1975) and the recent fisheries literature.

Example 9.8 Calculations for catch-effort models

There are a number of ways to estimate variances and confidence limits for the catch-effort models. These are summarized by Seber (1973,1982: Ch.7). Most current usage is based on regression calculations, where the models are represented as:

$$y = a + bx$$

where y is either catch per unit effort (eq.9.21) or log C(t) as in eq.9.23, while x is cumulative kill or cumulative effort. We then recall that simple linear regression can be viewed as a straight line passing through the means of the observations:

$$y - \bar{y} = b(x - \bar{x})$$

or,

$$y = (\bar{y} - b\bar{x}) + bx = a + bx$$

and that b is estimated by:

$$b = \frac{\Sigma(y_i - \bar{y})(x_i - \bar{x})}{\Sigma(x_i - \bar{x})^2}$$

In both of the equations (9.21 and 9.23) k is estimated by b, while a is either kN or $\log_e(kN)$. Confidence limits for k can be

calculated directly, as in the usual linear regression calculation for b. However, the regression intercept (a) estimates the product kN, or its logarithm, so that an approximation is required for a variance estimate of N, which is estimated from:

$$\hat{N} = \frac{a}{b} = \frac{kN}{k}$$

or by solving for N in:

$$\log_e (kN) = \bar{y} - \bar{x}$$

i.e.,

$$\hat{N} = \frac{\exp(\bar{y} - b\bar{x})}{b}$$

(students should remember that b will be negative in the present situation).

Seber (1973,1982:Ch.7) gives an approximate variance estimate for N appropriate for the logarithmic form (eq. 9.21) as:

$$V[\hat{N}] = s^2 \hat{N}^2 [\frac{1}{n} + \{\frac{b\bar{x} - 1}{b}\}^2 \{\frac{1}{\Sigma(x_i - \bar{x})^2}\}]$$

and for eq. 9.23 he gives:

$$V(\hat{N}) = \frac{s^2}{b^2}[\frac{1}{n} + \frac{(\hat{N} - \bar{x})^2}{\Sigma(x_i - \bar{x})^2}]$$

To use these equations in practice, we substitute the "variance about regression" for s^2. It will be recalled that this is estimated as the "mean square" of the deviations of the observation from the regression line, i.e.,

$$s^2 = \frac{\Sigma(y_i - a - bx_i)^2}{n - 2}$$

The numerator can be rearranged by using $a = \bar{y} - b\bar{x}$, grouping terms, and then writing out the squared terms:

$$\Sigma[(y_i - \bar{y}) + b(x_i - \bar{x})]^2 = \Sigma(y_i - \bar{y})^2 - 2b\Sigma(y_i - \bar{y})(x_i - \bar{x}) + b^2\Sigma(x_i - \bar{x})^2$$

Substituting the estimated form of b reduces this to:

$$\Sigma(y_i - \bar{y})^2 - b^2\Sigma(x_i - \bar{x})^2$$

and the two "sums of squares" on the right can be calculated from, for example,

$$\Sigma(y_i - \bar{y})^2 = \Sigma y_i^2 - \frac{(\Sigma y_i)^2}{n}$$

Variances of a and b are calculated as:

$$v(b) = \frac{s^2}{\Sigma(x_i - \bar{x})^2} \quad \text{and} \quad v(a) = \frac{s^2 \Sigma x_i^2}{n\Sigma(x_i - \bar{x})^2}$$

That for v(b) can be used directly to obtain approximate confidence limits on k, as:

$$k \pm t_{n-2}[v(b)]^{1/2}$$

Example 9.9 Variable-effort models

In the variable-effort models described above, the fishing mortality rate, F, was assumed constant. When this is an unrealistic assumption, an alternative is to consider mortality due to exploitation to be proportional to the effort, so that $F_i = kE_i$ where E_i is the effort in the i^{th} time period and k is a constant "catchability coefficient" (as used in eq.9.23 and 9.24).

In generalizing the model, one may as well also permit time intervals of varying length, so instead of a constant time period, t, we now use intervals $t_i - t_{i-1}$, where i = 1,2,3,..., and t_o denotes the beginning of exploitation. Thus where eq.9.25 contains a term s^{i-1} as the product of constant survival rates through the i-1 previous intervals, it now needs to be a product like $s_1 s_2 ... s_{i-1}$, which would then be replaced by:

$$s_i = e^{-(kE_i + X)(t_i - t_{i-1})}$$

and the equivalent of eq. 9.27 becomes:

$$kE_i = [\frac{kE_i N}{kE_i + X}][\exp(-S(E_j + X)(t_j - t_{j-1})][1 - \exp\{(kE_i + X)(t_i - t_{i-1})\}]$$

This new equation is not so readily treated by regression methods. One approach is to consider the ratio C_{i+1}/C_i, which, after taking logarithms (to base e) and rearranging gives a complicated equation that can be replaced by an approximate solution by dropping the second logarithmic term and rearranging the result so that a single linear regression equation (y=a+bx) results with:

$$y = \frac{1}{t_i - t_{i-1}} \log_e [\frac{C_i E_{i+1}}{C_{i+1} E_i}] \text{ , and a=X, b=k, and x=}E_i$$

Estimates of the slope (b) and intercept (a) provide approximate values of the unknown quantities X and k. An improved estimate is then obtained by substituting these trial values of X and k in the second logarithmic term, and then using the entire right-hand side of the equation as y in a new regression calculation. The resulting estimates of k and X can be again substituted and the regression calculated again to check whether the estimates change enough to justify another cycle. Readers familiar with the series expansion of e^{-X} will note that one might start with y values as:

$$y = \frac{1}{t_i - t_{i-1}} \log_e [\frac{C_i E_{i+1}[t_{i+1} - t_i]}{C_{i+1} E_i [t_{i+1} - t_{i-1}]}]$$

to give a somewhat improved first approximation.

A fault in the above procedure was noted by Chapman(1961), in that the successive y values will be correlated, again violating the essential assumptions for ordinary linear regression. To see this, one can examine the first 3 values of y:

$$y_1 = \frac{1}{t_1 - t_0} \log_e \frac{C_1 E_2}{C_2 E_1}$$

$$y_2 = \frac{1}{t_2 - t_1} \log_e \frac{C_2 E_3}{C_3 E_2}$$

$$y_3 = \frac{1}{t_3 - t_2} \log_e \frac{C_3 E_4}{C_4 E_3}$$

from which we see that successive terms contain common elements, e.g., E_2/C_2 appears in y_1 and again in y_2 (inverted). One way to avoid this is to restrict the regression calculations to every other data point; one could thus do two separate regression calculations and average the resulting estimates.

Evidently the above calculations will be somewhat involved and tedious. When effort varies markedly from period to period, an approach like this one seems to be the only realistic answer. Readers should note that when effort is constant from period to period, the basis for this procedure collapses. In fact, other things being equal, the wider the range of effort encompassed by the study, the more information one gets about k and X.

Example 9.10 Catch-effort data on an "open" population

Some data on a population of tagged juvenile cottontail rabbits may be used to study the effects of natural mortality (Eberhardt et al.1963). The data apply to the 1955-56 hunting season, and were selected from 7 years of similar data as giving the best fit to a regression line. Hence the variance about regression (or correlation) should not be regarded as typical of such data. In the first 3 weeks of hunting, 32 tagged individuals were harvested, but the effort data are not usable, since this period encompassed heavy hunting for pheasants.

Weeks of season	Tagged animals shot C_i	Effort in gun-hours E_t	y log(C/E)	x E(t)
4-6	23	984	9.151	0
7-9	22	1167	2.936	.984
10-11	13	1042	2.524	2.15
12-15	9	1059	2.140	9.193

A total of 230 tagged juvenile rabbits constitute the pre-season population. Regression calculations gave the line: $y = 9.196 - .3213x$

One would thus estimate the initial population size as:

$$\frac{kN}{k} = a/b = 9.196/.321 = 99$$

If we add in the 32 rabbits killed in the first 3 weeks ("pheasant season") this gives 131 rabbits as compared to the 230 tagged before hunting began.

Turning to eq. 9.28, we see that the intercept and slope now represent a more complicated expression:

$$a = \log\left[\frac{FN}{F+X}(1-\exp[-(F+X)]\ )\ \right]$$

$$b = (F+X)^t$$

From this we can estimate:

$$\frac{FN}{F+X} = 88.92.$$

Some other data yield an estimate of the instantaneous rate of natural mortality of 0.089 (calculated on a monthly basis). Using this rate and assuming the above periods to be uniformly two weeks long, and 4 weeks in a month, students should estimate N. To compare the outcome with the initial population (230), assume 4 weeks of natural mortality between tagging and the beginning of "pheasant season" (which was 3 weeks in length).

9.13 Catch-effort models with non-harvest losses

Sources of mortality have commonly been divided into two categories, one due to exploitation and the other due to other causes, operating concurrently with the harvest, and usually described as "natural" mortality. When tagged individuals are included in the analysis it seems best to refer to rates due to exploitation and to "other" causes in order to be able to provide for the additional mortality due to tagging, and the possible effects of loss of tags and non-reporting of tags.

The elementary model can be introduced by considering survival over some fixed period as a constant rate, s, and supposing that harvest takes a fixed fraction, f, of the population available during the interval. Letting the initial population size be N, we then have the first catch (C_1) as:

$$C_1 = f(1-s)N$$

which can be interpreted as saying that, of the proportion dying in the period, a fraction (f) are taken by the harvest method. The number surviving up to the next period is just sN (since (1-s)N died in the first period), so that:

$$C_2 = f(1-s)sN$$

and, since s^2N survive to the third period:

$$C_3 = f(1-s)s^2N$$

and, in general:

$$C_i = f(1-s)s^{i-1}N \qquad (9.25)$$

The notion of "competing risks" (see Ricker 1975), or instantaneous rates, leads to defining:

$$s = e^{-(F+x)t} \quad \text{and} \quad f = \frac{F}{F+X} \qquad (9.26)$$

where F denotes the rate of exploitation (F for fishing) and X the "other loss" rate. When the other losses are assumed due only to natural mortality, the symbol M is commonly used rather than X. Since the rates are "instantaneous" and thus are independent of length of the time period, it is necessary to define an arbitrary interval length, t, and write equation (9.25) as:

$$C_i = [\frac{FN}{F + X}] e^{-(F+X)t(i-1)} [1 - e^{-(F+X)t}] \qquad (9.27)$$

If t is now defined in terms of a "unit of effort" then equation (9.27) serves to replace eq.(9.22) in circumstances where the population is "open" to other sources of loss. In effect, units of effort are regarded as operating sequentially on the population, so that we have the effect of i time intervals, each of length t during which the losses take place. In practice, many units of effort are applied simultaneously to the population. The model may nonetheless be satisfactory, but the effect will be one of making estimates of F and X differ from year to year or place to place, depending on how the effort is applied (how it is distributed in time and space).

A number of schemes have been devised to estimate the parameters of eq.(9.26) from actual data. One of the simplest is to take logarithms (using natural logs) giving:

$$\log C_i = \log[\frac{FN}{F+X} \{1 - e^{-(F+X)t}\}] - (F+X)t(i-1) \qquad (9.28)$$

and letting $y = \log Ci$, $a = \log[\frac{FN}{F+X} \{1 - e^{-(F+X)t}\}]$, $b = (F+X)t$, and $x = i - 1$, which gives a simple linear regression model:

$$y = a + bx.$$

This can readily be fitted to data on catch per unit effort and time (or cumulative effort). Since t is presumably known, the regression slope (b) gives an estimate of F+X. and interest then centers on estimating F, and possibly N, from the regression intercept (a).

When a tagging or marking study is used, N is a known quantity and the regression intercept can be written as:

$$a = \log F + \log N + \log[\frac{1 - e^{-(F+X)t}}{F+X}] \qquad (9.29)$$

so that an estimate of log F can be obtained by subtracting log N and an estimate of the quantity in brackets calculated from the slope estimate (b) of (F+X) and t. An immediate problem with estimation from eq.(9.28) is that the conditions of such a study do not conform very well to the assumptions required by a simple linear regression estimate.

In many situations, it is unrealistic to regard F as being constant from time unit to time unit, because of substantial variations in the effort expended in each time unit. Although grouping of units of effort has been used to produce new units of roughly equal size, such a practice is not very satisfactory if X, the "other loss" rate, is constant over time.

9.14 Sampling for indices

The various uncertainties about interpretation of indices, combined with conditions of their use, do not make it any less important to conform to good sampling practice in the use of indices. By and large, most index methods are applied on rather large and heterogenous areas so there is generally a prospect for considerable gains in efficiency (and the concomitant reduction of effort required) through the use of standard methods, like stratification. Since indices are usually rather tightly tied to seasonal conditions, a sampling design needs to be arranged to permit study of seasonal effects, too.

Specific sampling methods need to be tailored to each index, and this is best done by writing a model of the kind already described but containing parameters relevant to the major factors that must be considered in the particular instance at hand. Where possible, it seems desirable to arrange the design in two or more stages. The first stage should be constructed in terms of those factors that are fairly well understood, so that the effects of uncertainties about a particular index are most likely to crop up in the second (or lower) stage. In nearly all cases, one might expect to do a reasonably efficient job of stratification by area and time, and to restrict the effect of other factors on an index to subsampling within the strata.

Stratification very likely will need to be more nearly of the kind generally described as "analytical" in view of the hazards of getting area effects entangled with strata. That is, generally one would make an individual stratum out of as many contiguous units as possible, and try to avoid having scattered units belonging to the same stratum. However, if calibration is attempted, one cannot usually have more than a few strata, or costs of the requisite independent density estimates become too high.

9.15 Transformations

Most of the few statistical analyses of index data done so far have employed transformations. The stated purpose of the transformation usually is to attempt to achieve normality. There are, however, other requirements for the analysis of data of this kind. One arrangement is:

(1) additivity
(2) constancy of error variance
(3) normality of distributions
(4) independence of observations.

The usual goal in making a transformation is to carry out an analysis of variance. Thus one might have results of, say, roadside counts carried out over several years and in a number of different areas, and wish to know if there are statistically significant differences in the underlying population levels between areas and among years. Nearly all of the published analyses of index data have been so handled. An example of the analysis of variance on index data appears in Example 9.4.

Securing independence of observations is largely dependent on how the observations are taken. Constancy of variance and normality of error distributions are often taken as one feature, while additivity implies that the underlying model is linear, that is, of the form:

$$x_{ij} = A_i + B_j + e_{ij} \qquad (9.30)$$

so that an index value in year i and area j depends on an effect due to years (A_i), one due to areas (B_j) and a "random error component" (e_{ij}) which should be normally distributed with $E(e_{ij}) = 0$ for analysis of variance purposes. The usual analysis of variance table gives components associated with the elements of eq. (9.30) and uses the estimate of $V(e_{ij})$ as a basis for tests of significance of the "contributions" of year and area terms to the total variability.

Most of the available analyses of ecological data seem to proceed along the lines of seeking a transformation that tends to normalize the data, and then assuming that eq.(9.30) is satisfied. This is a dangerous course, if the investigator is really interested in sorting out the various effects influencing his data. In that case, the essential need is to achieve linearity (additivity) in the model. From the form of equations (9.22), and (9.27) it is evident that a logarithmic transformation will often be required. If the true relationship of index and density is as postulated in these equations, taking logarithms effectively converts them to additive (linear) models like eq.(9.30). If the results are to conform fully with the analysis of variance requirements, then one must also assume that an error component is multiplicative, i.e., for a roadside count conducted in the same area for a number of years and areas one would write the model as:

$$x_{ij} = B_j D_{ij} e_{ij}$$

so that:

$$\log x_{ij} = \log B_j + \log D_{ij} + \log e_{ij} \qquad (9.31)$$

and investigate the prospect that there are area effects on the proportionality coefficient (B_j) as well as assessing population differences. If the error component is assumed to be multiplicative, and it is further assumed that taking a logarithm converts it to a normal distribution, then those two assumptions imply that the original distribution was lognormal in form. However, such assumptions may be somewhat fictitious as evidenced, for example, by the previous remark that pellet-group data follow the negative binomial distribution reasonably well. However, the difference between negative binomial and lognormal distributions may not have much of an effect after log transformations.

Use of the analysis of variance technique on index data calls for rather more detailed study than is possible here. The major point to be made is that blind use of a transformation seems extremely unwise. In most practical cases, the experienced investigator really has little interest in some of the tests of significance. He will have long since concluded that there are real differences in population density between areas, and will largely be concerned with year to year changes on individual areas, and with studying the magnitude of differences between areas. He will also want to look for effects due to observer, as might be identified in Example 9.3 in logarithmic form (where i pertains to observer and j to area):

$$\log X_{ij} = \log B_i + 2 \log r_j + \log D_j + \log e_{ij}$$

Analysis of some of the other equations might be considered in similar terms, but, as already noted, the usual approach is through regression methods. Regression equations were previously mentioned here as tools for estimating some parameter (e.g., population size) in the model, but they can also be studied in an analysis of variance format.

One further aspect of transformations is that investigators tend to be somewhat uncomfortable about presenting results of an analysis in terms of the transformed variable. For the roadside count example, one may originally have a variable expressed as so many animals seen per mile of driving (or per hundred miles, hours of observation, etc.). The logarithmic transformation yields an "unnatural" kind of datum. The usual advice is to transform back to the original scale, neglecting the prospect that such a transformation will introduce a bias. Sometimes this may be the best procedure, particularly if the results are to be used extensively for administrative guidance. However, it should be remembered that we are really considering the use of an index and there should not be any special disadvantage to using an index in logarithmic units. In fact, the behavior of the index (additivity) may be much more suitable on the transformed scale. Consequently, It is often not desirable to transform back to the original scale. Staying in the transformed scale is particularly desirable if several indices are combined, as described in the next section.

9.16 Combining indices

The problem of combining population indices has not been investigated in any detail in ecological studies as yet. Related problems exist in econometrics, but the results obtained there have not been applied in ecology as yet. We will assume that several independent indices are available, and that the need is to combine them to construct a single, overall index. Independence here means derived from different and unrelated sources, and thus refers largely to the sampling methods.

Part of the problem is that the individual indices will generally have different scales of measurement, in consequence of the kinds of information on which they are based. One possible structure for an index value is:

$$X_{1i} = \beta_1 D_i + e_{1i} \qquad (9.32)$$

where D_i represents the true density and β_1 represents a "scale factor" or a "proportionality constant", while e_{1i} represents a random error component. We thus assume that a given index value is proportional to the true density, but has associated with it a randomly selected fluctuation, due presumably to a variety of influences on the observational process. The usual approach in problems of this type is to assume that $E(e_{ij}) = 0$.

Equation 9.32 can be used to indicate what results when correlation or regression techniques are applied to indices. Supposing we have a second index, with structure:

$$X_{2i} = D_i + e_{2i}$$

Then an indication of the behavior of the correlation and regression coefficients, assuming large samples, can be given as follows:

$$E(s^2_{x_1}) = \beta_1^2 \sigma^2_{D_1} + \sigma^2_{e_1}$$

$$E(s^2_{x_1}) = \beta_2^2 \sigma^2_{D} + \sigma^2_{e_2}$$

$$E(s_{x1x2}) = \beta_1 \beta_2 \sigma_D^2$$

where σ_D^2 refers to the true variation in population density over the set of areas being investigated. Then the sample correlation coefficient (r) is approximately:

$$r = \frac{s_{x1x2}}{[s_{x1}^2 s_{x1}^2]^{1/2}} \qquad (9.33)$$

$$= \frac{\beta_1 \beta_2 \sigma_D^2}{[\{\beta_1^2\sigma_D^2 + \sigma_{e1}^2\}\{\beta_2^2\sigma_D^2 + \sigma_{e2}^2\}]^{1/2}}$$

so that r approaches unity only if σ_{e1}^2 and σ_{e2}^2 are nearly zero. Consequently, a very high correlation between two indices might be taken to mean that both are nearly directly proportional to true density. However, this depends on the indices being obtained from independent sources (otherwise the two indices may simply have highly correlated errors), and on the correctness of equation 9.32.

The regression coefficient (b) is approximately:

$$b = \frac{s_{x1x2}}{s_{x2}^2} = \frac{\beta_1\beta_2\sigma_D^2}{\beta_2^2\sigma_D^2 + \sigma_{e2}^2)} \qquad (9.34)$$

so that it is necessary for $\sigma_{e2}^2 = 0$ before the regression coefficient will reflect the true ratio of the two coefficients. This is the situation where one index is exactly proportional to true density, and one that is not likely to be encountered in practice. It is also the usual condition for regression analysis (i.e., that the independent variable be measured without error). If the ratio of σ_D^2 and σ_{e2}^2 is somehow known, then the regression coefficient can be estimated without bias. In the usual circumstances, one does not know the ratio, and the problem becomes difficult to handle.

In many cases, the investigator will have at least a rough idea of the sampling effort that went into each index. Such information can serve as a source of weights for combining indices -- usually just by converting the "sample sizes" to proportions and multiplying the transformed index values by these weights. Before doing so, it is necessary to convert the indices to the same scale. It also seems essential to have roughly the same spread of values for each index. One simple way to achieve this result is to transform the several indices to have about the same mean and variance. If the mean and variance are selected as some convenient values, say Z and S^2, then the coefficients A and B for the transformation of the i^{th} index:

$$Z_i = BX_i + A$$

are obtained by noting that the variance of the transformed index should be equal to the original index adjusted by a constant:

$$\frac{\Sigma(z - \bar{z})^2}{(n-1)} = S^2 = \frac{B^2\Sigma(x-\bar{x})^2}{(n-1)}$$

so that:

$$B = S/s, \tag{9.35}$$

and since $Z = B\bar{x} + A$ we have:

$$A = Z - (S/s)\bar{x}. \tag{9.36}$$

A new set of coefficients (A,B) has to be calculated for each set of index data. The transformed index values are then combined by weights obtained as suggested above:

$$Y = W_1z_1 + W_2z_2 + \ldots + W_nz_n \tag{9.37}$$

where there are n indices and the weights (W_i) sum to unity. A transformed value will, of course, be obtained for each area and time period under study, so that the y of eq.9.37 might be written as y_{ij} to pertain to the i^{th} area and j^{th} time period. Then the index values would have to be written as x_{ijk} (k for k^{th} index), and so on.

All of the above pertains to a model (eq.9.32) which assumes that the errors are additive. In Sec.9.12 it was pointed out that the logarithmic transformation implies that the model really is:

$$X_i = \beta D_i e_i$$

and the logarithmic transformation presumably makes it possible to investigate index behavior over a set of areas. Combining different indices will then partly be conditioned by the results of the statistical analysis of individual indices, and clearly can become quite complicated.

9.17 Converting indices

It has already been noted here that there are circumstances where an index value is not suitable for management purposes. This suggests a need for ways to convert an index to an estimate of actual density -- or for "calibration" of an index. If a set of areas exists on which true population density can be estimated, then an apparent solution is just to compute the regression of true densities on index values. However, it seldom is possible to measure true density without error -- ordinarily some sampling process is involved. This then puts us in exactly the same position as led to eq. 9.34, except that it may now be assumed that $\beta_1 = 1$, so we have:

$$b = \frac{1}{\beta_2 + \dfrac{\sigma^2_{e_2}}{\beta_2\sigma^2_D}} \tag{9.38}$$

and a knowledge of the ratio of the two variances is required to obtain an estimate of $1/\beta_2$ for conversion or calibration purposes. Some elements of strategy for planning calibration work are evident from the above relationship, though, i.e., keep σ^2_{e2} as small as possible and choose the set of population densities used to have as wide a range as possible. Very likely these

may be conflicting aims, since the choice of low densities may tend to increase $\sigma^2_{e_2}$.

When the regression relationship can be assumed to go through the origin, and if the variance of y increases proportionately with increasing x, then the ratio estimate (cf.Cochran 1977) is known to be optimum and unbiased. Using the same model as above, we have:

$$E(\Sigma y_i)/E(\Sigma x_i) = 1/\beta_2$$

so this estimate is approximately unbiased. In most real-world situations, these are the more likely assumptions than those of linear regression, so ratio methods are to be preferred to regressions, here.

As mentioned before, there are reasons to doubt the accuracy of the usual approximations to variance estimates (and hence confidence intervals) arising from the presence of errors in the independent variable (X). Very possibly there should be another component of variance in the equations. A related issue has to do with the distinction between ratio estimators (total of the X_i assumed known) and double sampling (only a sample of the X_i assumed known). With measurement errors in the X_i, one can have an observation on X_i on every study unit and still not "know" the total. That is, were a new survey possible, it would not give the same total.

A somewhat pessimistic view of what is known about double sampling with errors in the independent variable can be tempered, however, by another look at eq.9.33. As remarked there, the sample correlation coefficient (r) approaches unity only if measurement errors in X and Y are nearly zero. Hence observing sample correlations on the order of 0.9 or better with index data, gives one some considerable encouragement to think that the measurement errors are small, at least in relation to σ^2_D . Hence, there are some grounds to suppose that the bias in eq.9.38 may not be unreasonably great, supposing β_2 is not very small.

9.18 Comments on the use of index data

Although Fig. 9.1 suggests a variety of ways to approach the use of indices, a great deal of research may be required before really suitable methodology is available for dealing with indices. Some of the problems may be worth mention here. One very important issue is selecting a model for the analysis. Using a simple multiplicative model has several advantages. One is the log-transformation which tends to "normalize" the data. The second is that it produces a linear relationship with time, if the population is changing at a constant rate (e.g., growing exponentially). One very important question is whether prospective auxiliary variables do in fact have a multiplicative effect on the index, so that the log transform yields a simple linear model. Perhaps an even more important question is whether the actual population is changing at a constant rate.

At present, the main approach to dealing with a situation where a number of variables may be involved seems to be to identify any variables that may possibly be relevant (and, of course, that can be measured!). Because the

underlying relationship may not be linear, a squared term is sometimes introduced in the set of variables to be studied. An effort is then made to find out which of the several variables considered may be worth keeping in a final model. One might, for example, want to consider 5 candidate auxiliary variables. Including squared terms for each, then yields 10 variables. The immediate problem is one of reducing the set to some smaller number of "significant" variables. One approach is to compute all possible regressions with each of the candidate variables left out in turn, and with that variable included, and to use one or more criteria to see whether inclusion of a given variable is worthwhile. Three such criteria are frequently considered, one being the regression mean square, the second R^2 [eq.(9.11)] and the third Cp [eq.(9.17)}.

With as many as 10 variables, this approach gets out of hand, insasmuch as there are 2^r equations to study ($2^{10} = 1024$). An alternative is to start with an equation containing all 10 variables and use a pre-determined procedure to work back through the list and to eliminate those variables that do not meet certain criteria. A third procedure is stepwise regression, sarting with the "best" variable (most highly correlated with the index) and work through the set, adding a new variable if it meets a significance criterion, and stopping when there isn't a significant improvement.

Because a lot of computing is involved, these procedures generally depend on a computer program. Different results may be obtained from different procedures, and depending on the order in which variables are introduced, and opinions on how to proceed may differ from reference to reference. One should always examine (plot) the residuals from regression to see whether they provide any hints as to possible improvements.

The Cp criterion seems to me unlikely to be very useful for index studies. As described by Draper and Smith (1998), it may be most useful in situations where there is some reason to believe that most of the variables relevant to the process being studied are included in the set to be analyzed. Cp may then serve in picking out sets that are somehow "adequate" to describe the process. It is unlikely that one can hope to measure most of the variables that influence a population index.

The likely situation can be suggested for the bear data of Example 9.6. Fig. 9.20 shows the data of Fig. 9.4 with the regression mean square of the adjusted bear index [eq. (9.18)] added. The inclusion of auxiliary variables has reduced the regression mean square to about 40% of that of the unadjusted index, and thus presumably yields a better index of population trend. However, there are a lot of smaller regression mean squares in the data set of Fig. 9.4, presumably because these populations were more accurately enumerated. Using the delta method to aproximate the variability of the bear population (assuming survival rates and population sizes suggested by Eberhardt et al. (1994) and Eberhardt and Knight (1996) gives roughly the "true" variability shown at the far right side of the figure. Clearly a good deal of the variability in the index is unaccounted for. Sampling or measurement error in determining two of the auxiliary variables (frequency of sighting and snow depths) may account for some of the difference, but certainly not all.

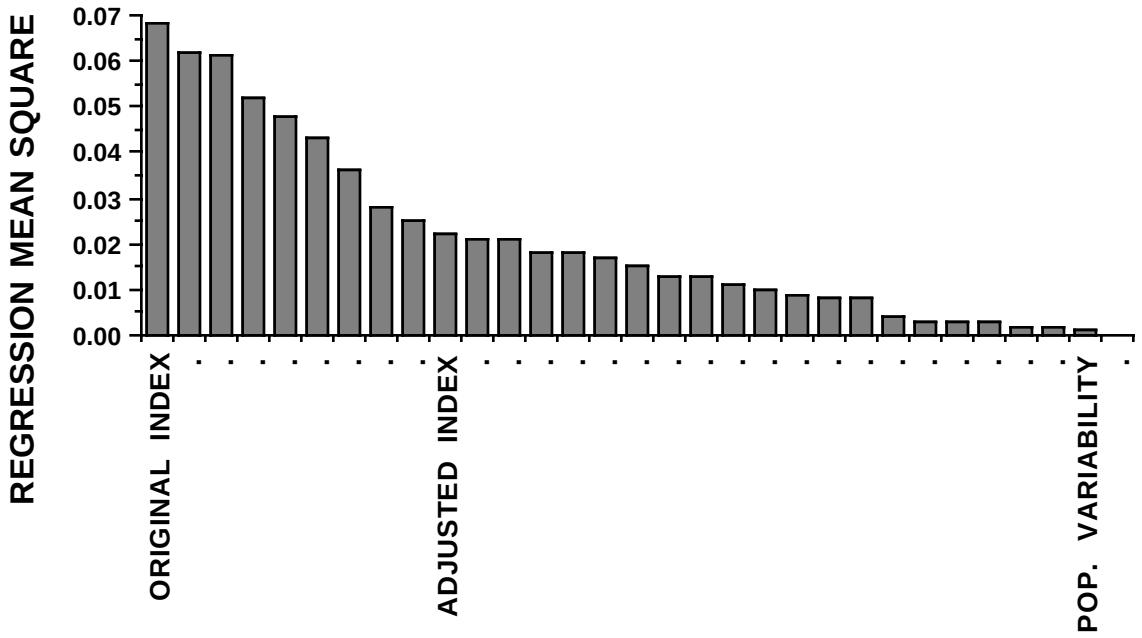

Fig. 9.20. Data of Fig. 9.4 with adjusted bear index and estimates of actual population variability added.

One may thus be inclined to doubt the utility of Cp in ecological index studies. Because it utilizes RSSp, the regression sum of squares, it will likely follow the trend indicated by the regression mean square. Also, RSSp for the model with all significant parameters included is the same as RSS(full) of R^2. Hence, it may be best to use regression mean square and R^2 as criteria for studying potential indices by multiple regression. One can interpret the resulting regression mean square by comparison to other values plotted in Fig. 9.4. It might also be noted that an earlier comment (Sec. 9.9) that "R^2 cannot become unity" is strictly true, but the very low value of estimated population variability for bears (Fig. 9.20) suggests that R^2 will become very close to unity when variability in regression mean squares is mainly from stochasticity. A number of such examples are present in the data used for Fig. 9.4, as shown in fig. 9.13.

9.19 Exercises.

9.19.1 Plot the data of Example 9.2 and suggest how one might go about trying to determine whether the two estimates of b given in the example are different.

9.19.2 Plot the data on Gambel quail from Example 9.3 and the ratio and regression lines. Which of the two lines appears to fit the data best? Why?

9.19.3 Conduct the analysis of variance described in Example 9.4

9.19.4 Data for counts of bison and for muxkox (Fig. 9.3) are given below. Using natural logarithms of the data fit a straight line and the “quadratic” of eq. (9.4) and use the F-test of eq. (9.6) to test for curvilinearity. Discuss the difference in results between bison and muskox. Plot the residuals to see if they help in explaining the difference.

YEAR	NO. OF MUSKOX	YEAR	NO. OF BISON
1950	61	1909	37
1951	76	1910	48
1952	77	1911	70
1953	90	1912	85
1954	100	1913	104
1955	116	1914	130
1956	126	1915	164
1957	143	1916	194
1958	181	1917	240
1959	206	1918	295
1960	256	1919	367
1961	293	1920	420
1962	353	1921	479
1963	406	1922	554
1964	467		

9.19.5 Using the data on brown bears of Table 9.3 without year 6 (only one point) conduct a one-way analysis of variance for differences among years using untransformed data and then using a log-transform. Which seems to be the best approach? Why?

9.19.6 Estimate lambda for the gray whale data of Table 9.4 using eqs. (9.7), (9.8), and (9.9). Compare the resulting estimates with the rate obtained from log-linear regression on the data.

9.19.7 Calculate the Durbin-Watson test [eq.(9.13)] on the residuals from log-linear regression from the gray whale data of Table 9.4. Report the value of d and comment on its meaning. What is the mathematical relationship between the slope and lambda for Table 9.7?

9.19.8 Repeat the calculations for Example 9.5 (sea otter census). Estimate the ratios for each day from eqns. (9.7), (9.8), and (9.9) and compare with the ratio estimates obtained by pooling the data from all 3 days. Do they suggest some differences? How might you test for significant differences between days?

9.19.9 Data for the counts of Yellowstone grizzly family groups (Example 9.6) appear below (note that the snowpack data have been "centered"—deviations from the mean are shown).. (1) Compute R^2 and s^2 for loglinear regression on the original counts. (2) Then compute R^2 and s^2 for the full data set, i.e., using the two auxiliary variables along with year. (3) Extend the computations to include t^2 (year-squared) as an auxiliary variable. Discuss your results. What is the effect of t^2?

LNCOUNT	YR	FREQ	SNOPAK
2.833	1	1.64	7.568
2.565	2	1.5	10.368
2.197	3	1.28	-18.432
2.565	4	1.08	3.868
2.485	5	1.4	1.568
2.639	6	1.58	-4.432

2.398	7	1.62	-9.832
2.565	8	1.2	9.568
2.833	9	2.29	2.868
2.197	10	2	-4.332
3.219	11	3.12	-0.632
2.565	12	1.64	3.568
2.944	13	2.12	-15.732
2.773	14	1.86	-3.132
3.219	15	1.95	8.968
3.178	16	2.65	-4.732
3.135	17	1.65	0.468
2.996	18	1.67	-6.332
2.996	19	1.47	3.268
2.833	20	1.47	-5.332
3.497	21	1.96	12.668
3.434	22	2.95	8.168

9.19.10 Do part (2) of Exercise 9.19.9 but use the actual years (1967, 1968, 1969,…) instead of 1,2.,3 and compare the results. Some references recommend "centering" the data. That is, instead of using 1967, 1968, 1969,… find the mean of this column and use the deviations from the mean as the x-variable. Try this and compare your results.

9.19.11 Note that (Fig. 9.16) the observed count for 1985 differs considerably from the predicted value. Do you think that value should be dropped from the index? Can you justify your answer statistically? How?

9.19.12 Compute Cp and AIC for the grizzly bear data using all of the auxiliary variables in the table. To compute AIC you need to know that the log-likelihood value for a linear model with normal errors is just the usual s^2., but calculated with n as divisor rather than n-p. That is, compute the sum-of-squares and divide by n. The assumption of normal errors is not strictly defensible statistically, but the log-transform seems to result in quite symmetrical distributions from regression given reasonably large samples.

9.19.13 Lobster catch data

DeLury (1947) gave the following catch data for lobsters:

Date	C(t)	K(t)	E(t)
May 23	82	0	0
24	75	7	8
25	94	13	16
26	80	16	19
27	83	22	27
29	89	25	32
30	70	32	40
31	58	37	48
June 1	64	40	53
2	55	45	61

5	52	50	69
6	45	53	76
7	45	54	77
8	49	55	79
9	45	57	85
10	48	60	90
12	43	62	96

Using eq.9.23 he estimated k = .008348 and N = 112.34, using simple linear regression as outlined above. From eq.9.24, he obtained k = .007984 and N = 116.33, also by linear regression. Students should repeat the calculations, and calculate variances and confidence limits.

9.19.14 Apply the expression for confidence limits given at the end of Example 9.5 to the 2 values of k obtaiend in the lobster catch data of Exercise 9.19.3. Report your results.

Exercise 9.19.15 Removal trapping

Zippin (1956) illustrated removal methods by assuming catches of 165, 101, and 54 animals were caught and removed in 3 nights of trapping. Calculations for eq. 9.21 are then based on the following data:

i	C_i	$y_i = \log_e(C_i)$	$x_i = i-1$
1	165	5.1060	0
2	101	4.6151	1
3	54	9.9890	2

Regression calculations will then proceed just as in the examples above, but we now have b = log(1-p). When p is small, we can represent log(1-p) by -p and the calculations are essentially those already described. However, in this case, p is clearly not small, so that eq. 9.21 is appropriate. Students should carry out the regression calculations and compare the estimate of N obtained by assuming b = p and b = log(1-p). Improved methods of estimation are available (Zippin 1958; Seber 1973,1982:Ch.7) but require an iteritive solution or the use of graphs (given in both references). Variance calculations are similarly complicated, and should be approached through the references cited.

Attention to variability is important in designing a study based on removal trapping, as a substantial fraction of the population must be caught in order to obtain a reasonably precise estimate. If we let q = 1-p, then the following equation (Zippin 1956:171) approximates the variance:

$$V[N] = \frac{N(1-q^n)q^n}{(1-q^n)^2\, q^{n-1}}$$

Students should try this equation for a few values of p, n, and N. Increasing n (beyond 3) doesn't do much to reduce the variance, which means that p has to be increased (by using more traps). This has a considerable practical

significance, in that it usually isn't sensible to run a removal trapping program very long, since immigrants will soon show up, violating the assumption of a closed population. Tagging some animals before the study starts is always a wise precaution, if feasible.

Exercise 9.19.16 Combining and comparing indices

In section 9.16 it was suggested that several different indices might be combined by: (1) transforming the individual indices to a common scale, and (2) weighting the transformed values by some independent measure of their variability. An example was given by Eberhardt (1960), portions of which are reproduced here to exhibit calculations. Four individual indices were used: (1) Accidental highway kills (recorded by Conservation Department staff. (2) "Camp kill" (estimates of deer taken on a special "camp" license, obtained through a mail survey of hunters). (3) July deer counts (a roadside tally by Conservation Department personnel. (4) Archery kill (deer taken on a special "bow and arrow" hunting license, estimated from an independent mail survey of those hunters).

An arbitrary transformation to a variance of 9.0 and mean (Z_i) of 4.0 was used as given by eqs.9.35 and 9.36. An example of the calculations appears in the second table below. Students should check their understanding of the equations by repeating the calculations. The transformed data were then combined into a single index as in eq.4.15, by using a set of weights, W_i that summed to unity. One possible choice of weights was the sample sizes for the various indices:

Index	Units used	Average number per year	Square root	Weight
July count	Number of deer seen	5000	70	.409
Archery kill	Number of hunters in samples	2000	44.7	.258
Camp kill	Number of parties in samples	200	14.1	.081
Highway kill	Number of deer tallied	1900	49.6	.252
				1.0000

Transformation of roadside deer counts to standardized values

	Deer seen per 100 hours						
District	1952	1953	1954	1955	1956	1957	1958
1	18.4	16.4	29.5	19.3	20.7	17.6	17.1
2	29.9	29.4	34.8	29.5	37.0	30.9	40.3
3	30.9	35.1	37.1	35.2	39.9	32.8	35.5
4	27.8	27.4	42.5	26.5	37.0	30.1	30.5
5	49.6	32.8	28.1	30.4	29.8	32.9	35.0
6	39.7	11.3	16.6	8.5	19.7	11.5	19.2
7	71.4	78.6	77.8	58.2	38.1	49.2	59.7

8	20.6	19.1	14.2	14.0	18.9	16.9	24.8
9	15.0	12.2	19.2	9.9	16.8	22.1	16.7

Transformed values

1	7.93	7.54	4.09	8.10	8.38	7.77	7.68
2	9.00	4.07	11.12	4.09	11.55	4.36	12.19
3	4.36	11.18	11.57	11.20	4.95	4.73	11.26
4	9.76	9.68	12.62	9.50	11.55	4.21	4.28
5	12.83	4.73	9.82	4.26	4.15	4.75	11.16
6	12.08	6.54	7.58	6.00	7.01	6.58	6.92
7	18.25	19.65	19.49	15.68	11.76	12.76	15.97
8	8.36	6.90	7.11	7.07	8.02	7.64	9.17
9	7.27	6.72	6.92	6.27	7.62	8.65	7.60

Data for transformation

s^2 =237.315 n=63
s=15.405 B=S/s=9.000/15.405=.1947
x=29.041 A=10-.1947(29.041)=4.3457

Check on transformation

s^2=8.9991
z=10.0002

Several checks on behavior of these weights were used, including: (1) error mean square in an analysis of variance of the index (compared as coefficients of variation), (2) mean square deviation from regression (against pellet count data), and, (3) correlations among the 4 indices. Since none of these analyses provided a measure that would be independent of the index observations, they were only used to check on behavior of the weights, i.e., to show that the weights were roughly correct.

Execution of a transformation to the same scale and appropriate weighting does not provide much evidence about validity of the combined index. In the present example, there were two independent measures that could be used for this purpose. One was the pellet group counts (already described here), and the other a population estimate based on sex composition, age structure, and kill (harvest) data, labelled the S-A-K method. Correlations between the 3 sources were:

	Combined index	S-A-K
S-A-K	.934	--
Pellet counts	.954	.951

There is thus independent evidence that the index did indeed provide a good measure of population levels.

In a Wisconsin study, McCaffrey (1976) introduced another index, a count of deer trails intersecting 0.4 km transects. Correlations were:

.	Trail index	S-A-K
S-A-K	.94	--
Pellet counts	.89	.781

It thus seems that this index is also well-correlated with independent measures of deer population size.

The question of how one uses a "good" index once it has be analyzed and shown to be well-correlated with direct measures of abundance is one needing further research. In broad outline, it has been suggested here that, when the direct estimate is available for only portions of the study area, while the index is available for the entire area, then one might use double sampling (ratio method) to use the index to arrive at an estimate for the entire study area. Presumably predictions of population levels might be made for sub-areas from the same relationship. These will, however, be quite variable, and just how confidence limits on these predictions might be obtained seems to me to be an unresolved question as yet. In areas where both index and direct measure are available, it does not seem feasible to use the index, unless it is converted to an estimate on the basis of prior (not current) experience. The current estimates should, however, be used in checking to see that the index is still "in calibration".